2022-12-01

Metrological Challenges Of Practical Computer-Enhanced Measurements

Hector Alejandro Reyes
University of Texas at El Paso

METROLOGICAL CHALLENGES OF PRACTICAL COMPUTER-ENHANCED MEASUREMENTS

HECTOR ALEJANDRO REYES

Master's Program in Computer Science

APPROVED:

Vladik Kreinovich, Ph.D., Chair

Shirley Moore, Ph.D.

Paras Mandal, Ph.D.

Stephen L. Crites, Ph.D.
Dean of the Graduate School

METROLOGICAL CHALLENGES OF PRACTICAL COMPUTER-ENHANCED MEASUREMENTS

by

HECTOR ALEJANDRO REYES, B.Sc.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Computer Science

THE UNIVERSITY OF TEXAS AT EL PASO

December 2022

Acknowledgments

I would like to thank all my committee members for their support:

- I want to thank my advisor Dr. Vladik Kreinovich for his supervision and advice,
- I want to thank Dr. Shirley Moore for informing me about many practical problems — some of which started this thesis, and for helping me better understand the related challenges, and
- last but not the least, I want to thank Dr. Paras Mandal for his detailed advice and help on how to describe my findings in a more professional manner.

I want to also thank the faculty of Computer Science Department, especially those whose classes I attended: Dr. Martine Ceberio, Dr. Yoonsik Cheon, Dr. Eric Freudenthal, Dr. Olac Fuentes, Dr. Mahmud Hossain, Dr. Luc Longpré, Dr. Daniel Majia, Dr. Oscar Mondragon, Dr. Saeid Tizpaz-Niari, Dr. Deepak Tosh, and Dr. Natalia Villanueva Rosales. The topics that I learned from them are invaluable lessons that I will continue to use throughout my career.

Most importantly, I would like to thank my friends and family. Without the support of my parents, my sisters, and my brother, I would not be where I am today. To my parents: Alejandro and Patricia who have always been there with me throughout my entire list, I would like to thank them the most. Without all the sacrifices that they made for me and my sistsers, none of us would be where we are. From making the decision to move to El Paso all the way to not just supporting but encouraging my decision to study computer science. They truly are the reason that I am the person that I am. Thanks to their immense and continuous support I have been able to reach heights that I previously did not believe I could achieve. Though they faced hardship, the unending peristence to continue moving forward has always inspired me to keep pushing even when I feel that I have nothing left

to give. To my siblings: Liz, Ale, and Jeandell I want to thank for always setting the bar high. Seeing the hard work that they put in to reach their dreams was always the second wind that I needed to push myself when I would have given up. Though they faced their own unique challenges, the grace and elegance with which they would face and overcome every challenge the world threw at them continues to inspire me. To the many friends that I made, lost, and reconnected with: I thank each and every one of you for being there for me through the many sleepless nights. For listening to me talk about topics that they did not understand but always showed genuine enthusiasm for. While there are too many friends throughout my studies to list, each and every one is as important as the last. Each of which has made a unique and profound impact in me, and each of which I will forever be grateful for.

Abstract

As technology progresses, sensors and computers become cheaper, so we can afford to perform more measurements and process the data faster. However, this also brings challenges. The goal of this thesis is to enumerate these challenges and to provide possible solutions.

The first challenge is related to the fact that the existing metrological recommendations are mostly based on the previous practice, when we could only afford to have a small number of measurements. In this regard, our objective is to describe the related problem and to propose a solution to this problem. These description (on the example of the design of the Thermonuclear Research Center) and proposed solution form the first contribution of this thesis.

The second challenge is related to the fact that in the past, when there were few affordable measuring instruments and we could only afford a few measurements, there were not that many options. So, we could select one of these options "by hand". Nowadays, with a potential to perform a large number of measurements and the availability of many different measuring instruments, the number of possible measurement options become large. Our related objective is to develop methods for optimal planning. Our related contribution is developing such a method for an important case of distributed measurements.

The third challenge is related to the fact that with the possibility to perform numerous measurements and process their results, we often encounter situations when for different pairs of measurement errors we have different types of information: some are known to be independent, for others, we do not have such information. Our objective is develop algorithms for dealing with such situations. Our contribution is to develop algorithms for the case when we have a small number of pairs with different type of information.

The final challenge is to extract useful information from all these measurement results. This extraction is the fourth objective of this thesis. Our contribution is in handling an important particular case of this objective: finding faults in a smart electric grid.

Table of Contents

Chapter 1

Introduction

1.1 Background: Practical Computer-Enhanced Measurements

All the information about the world comes from measurements – and from computer-based processing of these measurements; see, e.g., [20]. From the measurement viewpoint, the corresponding process can be divided into the following stages [20]:

- First, to make sure that the measurement results provide useful and reliable information, we need to set up some general principles about measurements – how to gauge the accuracy of measuring instruments, how to calibrate these instrument, etc. This information is usually codified in measurement-related ("metrological") standards and other documents.
- Second, we need to plan the measurements – and perform them.
- After that, we need to process measurement results to come up with useful information about the world.
  - In some cases, we already have efficient data processing algorithms. In such situations, from the metrological viewpoint, the main challenge is to understand how accurate are the results of data processing – i.e., how the measurement errors affect the result of data processing.
  - In many other cases, we do not yet have efficient data processing algorithms. In such cases, we need to come up with such algorithms.

1.2 Challenges Related to Practical Computer-Enhanced Measurements: What Is Known, and What Are Remaining Problems

As technology progresses, sensors and computers become cheaper. As a result, we can afford more measurements, and we can afford to process them faster and better. However, on all above-described stages, this progress also brings challenges. The overall objective of this thesis is to provide solutions to at least the simplest particular cases of these challenges.

The first challenge. The first challenge is related to the fact that the existing metrological recommendations are mostly based on the previous practice, when we could only afford to have a small number of measurements. As a result, the same system that in the past (when fewer measurements were possible) would have successfully passed the metrological analysis is no longer certified when more measurement results are available. This is a serious problem that, e.g., halted the design of the International Thermonuclear Experimental Reactor ITER; see, e.g., [9, 12].

The second challenge. The second challenge is related to the fact that in the past, when there were few affordable measuring instruments and we could only afford a few measurements, there were not that many options. In such cases, planning measurements simply meant selecting one of these options. So, we could plan the measurements "by hand". Nowadays, with a potential to perform a large number of measurements and the availability of many different measuring instruments, the number of possible measurement options becomes so large that we need to develop methods for optimal planning. There exist techniques for such planning – see, e.g., [3, 5, 6, 8, 11, 14, 18, 21, 22] – but these techniques are mostly based on limited number of measurements. For situations when we have a large number of measurements, to the best of our knowledge, no practical general methods are known – even for the simplest case when the data processing algorithms consists of simply

adding or averaging the measurement results.

The third challenge. The third challenge is related to the fact that in the past, when we could only afford a few measurements, these measurements were usually performed by similar measuring instruments, instruments for which we had a good understanding of what causes their measurement errors; see, e.g., [20]. In some situations, most measurement errors were causes by internal features of the instruments. In this case, the corresponding measurement errors were independent. In other situations, mostly external features were dominant, in which case we do have any information about the relation between different measurement errors. In both types of situations, formulas were developed for processing the resulting uncertainty. With the possibility to perform numerous measurements and process their results, we often encounter situations when some pairs of measurement errors are independent but for other pairs of measurement errors, we do not have any information about their relation.

The fourth challenge. The final – fourth – challenge is how to extract useful information from all these measurement results [20].

1.3 Specific Objectives of This Thesis

The main objective of this study is to deal with these four challenges – at least with the simplest possible cases of these challenges.

The first objective. Our first objective – related to the first challenge – is to explain how to makes sure that the measurement standards do not lead to the current counterintuitive practice of reducing the number of measurements.

The second objective. Our second objective – related to the second challenge – is to come up with optimal experiment design for the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results.

The third objective. Our third objective – related to the third challenge – is to come up

with techniques for processing measurement results in situations which are slightly different from the above-described well-studied ones; namely:

- for the situations when for most pairs of measuring instruments, we know that the corresponding measuring errors are independent, but for a few pairs, we do not have any information about their dependence, and
- for the situations in which for most pairs of measuring instruments, we have no information about the dependence between the corresponding measurement errors, but for some pairs, we know that the corresponding measurement errors are independent.

The fourth objective. Our fourth objective – related to the fourth challenge – is to extract information from the measurements, in the simplest case when we only know the ordering of the measurement results, but not the actual numerical values.

1.4 How We Are Contributing to These Objectives

Our contribution to the first objective. For the first objective, we propose the idea of how to change the standards, so as to avoid the above-mentioned unfortunate situations, when additional measurements can (and do) put the system at risk of not being approved.

Our contribution to the second objective. For the second objective, we provide a theoretical analysis of the problem and find a new explicit formulas for the optimal measurement design. As an interesting side effect of this theoretical analysis, we come up with an explanation of why measurement accuracy is usually described by listing absolute and relative error components. To the best of our knowledge, ours is the first theoretical explanation for this widely used practice.

Our contribution to the third objective. For the third challenge, we provide new explicit easy-to-implement formulas describing the uncertainty of the result of data processing in above-described situations.

Our contribution to the fourth objective. Finally, for the fourth challenge, we provide a theoretical result explaining – on the example of fault location in an electric grid – that information about the ordering of measurement results can be sufficient to accurately locate the fault.

1.5 Limitations of the Study and Remaining Problems

In all four cases, in this thesis, we only deal with the simplest possible cases of the general challenges.

Limitations and remaining problems related to the first challenge. For the first challenge – related to measurement-related certification of systems – we simply propose an idea, it is still necessary to develop this idea and to come up with the corresponding standards.

Limitations and remaining problems related to the second challenge. For the second challenge – related to measurement design – we only deal with the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results. It is necessary to extend our analysis to more complex data processing algorithms.

Limitations and remaining problems related to the third challenge. For the third challenge – of uncertainty analysis in situations when we have different information about different pairs of measurements – we only deal with the cases when for the most pairs, we have information of the same type, and only for a small number of pairs, we have different information. It is necessary to extend our analysis to situations when we have a larger number of pairs with different information.

Limitations and remaining problems related to the fourth challenge. Finally, for the fourth challenge – related to processing measurement results – we only deal with the case when we know the ordering of the measurement results, but not the numerical values

themselves. It is necessary to extend our analysis to situations when we have (and can use) numerical values as well.

1.6 Organization of the Thesis

We deal with our four objectives, correspondingly, in Chapters 2 through 5:

- Chapter 2 deals with the first objective,
- Chapter 3 deals with the second objective,
- Chapter 4 deals with the third objective, and
- Chapter 5 deals with the fourth objective.

The final Chapter 6 contains conclusions and recommendations for future work.

Chapter 2

Over-Measurement Paradox: Suspension of Thermonuclear Research Center and Need to Update Standards

In this chapter, we deal with the first of the four challenges outlined in Chapter 1. This challenge is related to the fact that

- while in general, the more measurements we perform, the more information we gain about the system and thus, the more adequate decisions we will be able to make,
- in situations when we perform measurements to check for safety, the situation is sometimes opposite: the more additional measurements we perform beyond what is required, the worse the decisions will be: namely, the higher the chance that a perfectly safe system will be erroneously classified as unsafe and therefore, unnecessary additional features will be added to the system design.

As we have mentioned, this is not just a theoretical possibility: exactly this phenomenon is one of the reasons why the construction of a world-wide thermonuclear research center has been suspended. In this chapter, we show that the reason for this paradox is in the way the safety standards are formulated now – what was a right formulation when sensors were much more expensive is no longer adequate now when sensors and measurements are much

cheaper. We also propose how to modify the safety standards so as to avoid this paradox and make sure that additional measurements always lead to better solutions.

2.1 What Is Over-Measurement Paradox

General case: the more measurements, the better. Most of our knowledge about the world comes from measurements; see, e.g., [20]. Each measurement provides is with an additional information about the world – and once we have a sufficient number of measurements of the same system, we may be able to find the equations that describe the dynamics of this system and thus, to get even more additional information that was hidden in the original measurements.

The more measurements we perform, the more information we gain about the system, the more accurate our estimates, and thus, the better will be our decisions. From this viewpoint:

- the more measurements we perform,
- the better.

We only expect one limitation on the number of measurements – the financial one. Indeed, at some point, after we have performed a large number of measurements, we get a very accurate picture of the measured system. Decisions based on this picture are close to optimal, and a very small expected increase in optimality may not be worth spending money on additional measurements.

Over-measurement paradox: case study. Most of our energy comes from the Sun. In the Sun, as in most stars, energy is generated by the thermonuclear synthesis, when protons – i.e., nuclei of Hydrogen (H) – combine together to form nuclei of Helium (He). This is a very efficient way of generating energy, a way that does not lead to pollution or other side effects. The majority of physicists believe that this is a way to get energy for our civilization: instead of relying on direct or indirect energy from the thermonuclear reaction

inside the Sun, why not use the same reactions ourselves – this will be a very effective and clean idea.

The idea is theoretically feasible, but technically, this is a very difficult task. Researchers and engineers all over the world have been working on it since the 1950s. To speed up the process, researchers from 35 major world countries decided to join efforts, and allocated \$65 billion dollars to build an international research center where specialists from all the world will work on this topic. This project is named ITER – this is both:

- an abbreviation of International Thermonuclear Experimental Reactor and
- the Latin word meaning "the way"; see, e.g., [9].

The problem is that as of now, this project is suspended, and one of the main reasons for this suspension is over-measurement; see, e.g., [12]. In a nutshell, the requirement was that, to guarantee safety, the level of danger – e.g., the level of radiation – was supposed to be below the safety threshold at a certain number of locations and scenarios.

- The current design does satisfy this criterion.
- However, the designers decided to be thorough and simulated more measurement situations.

Unfortunately, some of the expected measurement results exceed the threshold. As a result, the whole project is in suspension. Making sure that all future measurements satisfy the criterion would require a drastic redesign and a drastic further increase in the cost of the whole project – so drastic that it is doubtful that this additional funding will appear, especially in the current economic situation.

Why is it a paradox? If the designers did not perform these additional measurements, the design would have been approved and the project would have started. So in this case, additional measurements made the situation much worse – not only for the researchers, but for the humankind as a whole. This is a clear situation where additional measurements do not help at all.

But is it really a paradox? Maybe it is good that the project stopped – maybe additional measurements revealed that the original design was unsafe?

What we do in this chapter. In this chapter, we analyze the situation from the general measurement viewpoint and come up with several conclusions.

- first, we show that this situation is, in principle, ubiquitous: a similar problem will surface in many other projects, including those that have already been approved and designed and seem to function OK;
- second, although it may look that the problem is caused by insufficient safety of the original design, we show that this is not the case: practically any design, no matter how safe, will fail the currently used criteria if we perform sufficiently many measurements;
- finally, we propose a natural suggestion on now to change the corresponding standards so as to avoid such unfortunate situations.

2.2 Analysis of the Problem

Let us formulate the situation in precise terms. We are interested in studying states of different systems. A usual way to describe each state is by describing the values of the corresponding quantities at different locations and at different moments of time.

Usually, specifications include constraints on the values of some of these quantities. These may be constrains on the radioactivity level, constraints on concentration of potentially harmless chemicals, on the temperature, etc.

In all these cases, a typical constraint is that the value of some quantity q should not exceed some threshold q_0: $q \le q_0$.

How can we check this constraint: seemingly natural idea. In the ideal world, we should be able to measure the value $q(x,t)$ at all possible spatial locations x and for all possible moments of time t, and check that all these values do not exceed q_0.

Of course, in real life, we can only perform finitely many measurements. So, a seemingly natural idea is to perform several measurements, and to check that all measurement results $q_1, \ldots, q_n$ do not exceed q_0. However, it is known that this seemingly natural idea can lead to dangerous consequences; see, e.g., [20]. Let us explain why.

Why the above seemingly natural idea is dangerous. The actual value of the quantity q depends on many factors which are beyond our control. For example, the actual radioactivity level at a given location is affected by the natural radioactivity level at this location – the level that can change based, e.g., on weather conditions, when wind brings matter from neighboring areas where this natural level is somewhat higher. There are many small independent factors affecting the actual value of the quantity q.

In addition, the measurement result is somewhat different from the actual value of the measured quantity; see, e.g., [20]. We may be able to get rid of major sources of such measurement errors, but there are always a lot of small independent factors that lead to small changes of the measurement results.

Because of both types of random factors, the measured value differs from its locally-average level, and this difference is the result of a joint effort of a large number of small independent factors. It is known (see, e.g., [23]) that such a joint effect is usually well described by a normal (Gaussian) distribution. To be more precise:

- What is known is that in the limit, when the number N of small independent random factors increases (and the size of each factor appropriately decreases), the probability distribution of the joint effect of all these factors tends to the normal distribution – which thus appears as the limit of the actual distributions when N increases.
- By definition of the limit, this means exactly that when the number N of factors is large – and in many practical situations it is large – the actual distribution is very close to normal.

So, with high accuracy, we can safely assume that this distribution is normal.

This assumption explains why the above seemingly natural idea is dangerous. Indeed, what we have is several measurement results $q_1, \ldots, q_n$, i.e., in effect, several samples from the normal distribution. Usually, measurement errors corresponding to different measurements are practically independent – and the same can be said about the random factors affecting the value of the quantity q at different spatial locations and at different moments of time. From this viewpoint, what we observe are n independent samples from a normal distribution.

If we only require that $q_i \le q_0$, we thus require that $\max(q_1, \ldots, q_n) \le q_0$. Usually, our resources are limited, so we try to make the minimal effort to satisfy the requirements. In other words, when we institute more and more efficient filters – thus slowly decreasing the value q_i – and finally, reach the condition $\max(q_1, \ldots, q_n) \le q_0$, we stop and declare this design to be safe.

- We start with the design for which $\max(q_1, \ldots, q_n) > q_0$.
- So the first time when we satisfy the desired constraint $\max(q_1, \ldots, q_n) \le q_0$ is when we get

$$\max(q_1, \ldots, q_n) = q_0.$$

This again may sound reasonable, but it is known that the probability that the next random variable will exceed the maximum $\max(q_1, \ldots, q_n)$ is proportional to $1/(n+1)$. So:

- even if we perform 40 measurements – and this is, e.g., what measurement theory requires for a thorough analysis of a measuring instrument (see, e.g., [20]),
- we get a $1/40 \approx 2.5\%$ probability that next time, we will go beyond the safety threshold.

This is clearly *not* an acceptable level of safety – especially when we talk about serious, potentially deadly dangers like radioactivity or dangerous chemicals.

So what can be done to avoid this danger. To simplify our analysis, let us suppose that the mean value of q is 0. This can always be achieved if we simply subtract the

actual mean value from all the measurements result, i.e., for example, consider not the actual radioactivity level, but the excess radioactivity over the average value of the natural radioactivity background.

In this case, measurement results $q_1, \ldots, q_n$ form a sample from a normal distribution with 0 mean and some standard deviation σ.

- Of course, no matter how small σ, the normally distributed random variable always has a non-zero probability of being as large as possible – since the probability density function of a normal distribution is always positive, and never reaches 0.
- So, we cannot absolutely guarantee that all future values of q will be smaller than or equal to q_0.
- We can only guarantee that the probability of this happening is smaller than some given probability p_0, i.e., that

$$\text{Prob}(q > q_0) \le p_0.$$

So, to drastically decrease the probability of a possible disaster – from the unsafe 2.5% to the much smaller safety level $p_0 \ll 2.5\%$:

- instead of the original threshold q_0,
- we select a smaller threshold $\widetilde{q}_0 < q_0$ that guarantees that the conditional probability of exceeding q_0 is small:

$$\text{Prob}(q > q_0 \mid \max(q_1, \ldots, q_n) \le \widetilde{q}_0) \le p_0.$$

In this case:

- if we have n measurement $q_1, \ldots, q_n$ all below $\widetilde{q}_0$,
- then we guarantee, with almost-1 probability $1 - p_0$, that the next value will be below the actual danger threshold q_0.

This value $\widetilde{q}_0$ depends on q_0 and on the number of measurements n:

- the larger n,
- the larger the value $\widetilde{q}_0$.

When n increases, this value tends to q_0.

So what is included in the safety standard. When safety standards are designed, one of the objectives is to make them easy to follow:

- We do not want practitioners – who need to follow these standards – to perform complex computations of conditional probabilities.
- We need to give them clear simple recommendations.

From this viewpoint, the easiest to check if whether the measurement result satisfies a given inequality.

So, a reasonable way to set up the corresponding standard is to set up:

- the new threshold $\widetilde{q}_0$ and
- the minimal necessary number of measurements n.

The standard then says that:

- if we perform n measurements, and the results $q_1, \ldots, q_n$ of all these n measurements do not exceed this threshold $\widetilde{q}_0$, then the situation is safe;
- otherwise, the situation is not safe, and additional measures need to be undertaken to make this situation safer.

Resulting common misunderstanding. The fact that safety standards provide such a simplified description – and rarely mention actual threshold $q_0 > \widetilde{q}_0$ – makes most people assume that the critical value $\widetilde{q}_0$ provided by a standard is the actual danger level, so

any situation in which a measured value exceeds $\widetilde{q}_0$ is unacceptable. This is exactly what happened in the above case study.

And this is wrong conclusion:

- if we perform a sufficiently large number of measurements,
- we will eventually get beyond any threshold.

Indeed, according to the extreme value theory (see, e.g., [1, 2, 4, 7, 15]), for normal distribution with mean 0 and standard deviation σ, the average value A_n of the maximum $\max(q_1, \ldots, q_n)$ grows with n as

$$A_n \sim \gamma \cdot \sqrt{2 \ln(n)} \cdot \sigma,$$

where $\gamma \approx 0.5772$ is the Euler's constant

$$\gamma \stackrel{\text{def}}{=} \lim_{n \to \infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln(n) \right).$$

So, this mean value indeed grows with n.

Why this problem surfaces only now? Gaussian distribution was invented by Gauss in the early 19th century, measurements have been performed since antiquity, so why is this problem surfacing only now? Why did not it surface much earlier?

The main reason, in our opinion, is that, until recently:

- sensors were reasonably expensive – especially accurate ones – and the cost of measurements was non-negligible;
- in this case, while in principle, it was possible to perform more measurement than required for safety testing, this would have led to useless costs.

Lately, however:

- sensors have become very cheap: kids buy them to make robots, the cheapest cell phones have very accurate sensors of positions, acceleration, etc.;

- as a result, it is reasonably inexpensive to perform many more measurements than required;
- and, as we have mentioned, as a result, in situations that would previously – based on only the required number of measurements – would be classified as safe, now we get values exceeding the threshold $\widetilde{q}_0$ provided by the standard – and thus, we end up classifying perfectly safe situations as unsafe.

2.3 So What Do We Propose

What is the problem now: summarizing our findings. The reason why we have the over-measurement paradox is that current safety standards usually list only two numbers:

- the recommended threshold $\widetilde{q}_0$ and
- the recommended number of measurements n.

The idea that the results of all the measurements must not exceed $\widetilde{q}_0$ for the situation to be considered safe.

The problem is that the recommended threshold $\widetilde{q}_0$ is actually *not* the safety threshold q_0, it is smaller than the safety threshold – smaller so that for the prescribed number of measurements n, we would guarantee that:

- for all future values,
- the probability to exceed the real safety threshold q_0 should be smaller than the desired small value p_0.

When, in an actually safe situation, in which the probability to exceed q_0 does not exceed p_0, we perform more measurements than recommended, then it is eventually inevitable that some of them will be larger than the recommended threshold $\widetilde{q}_0$ – even though they will still, with almost-1 probability, be not larger than the actual danger threshold q_0. This leads to the following natural solution to the over-measurement problem.

Proposed solution: we need to change the standards. In addition to providing the two numbers $\widetilde{q}_0$ and n, we should provide the formula describing the dependence of the testing safety threshold $t(n')$ for different numbers $n' \geq n$ of actual measurements, so that for all n', we should have

$$\text{Prob}(q > q_0 \,|\, \max(q_1, \ldots, q_{n'}) \leq t(n')) \leq p_0.$$

At least we should provide the value $t(n')$ for several different values n', thus taking care of the cases when, due to thoroughness, practitioners will provide more measurements.

2.4 Conclusions and Recommendations for Future Work

Conclusions. In this chapter, we deal with the first of the four challenges of practical computer-enhanced measurements. This challenge is related to the fact that the existing metrological recommendations are mostly based on the previous practice, when we could only afford to have a small number of measurements. As a result, the same system that in the past (when fewer measurements were possible) would have successfully passed the metrological analysis is no longer certified when more measurement results are available. This is a serious problem that, e.g., halted the design of the International Thermonuclear Experimental Reactor ITER; see, e.g., [9, 12].

In this chapter, we propose the idea of how to change the standards, so as to avoid the above-mentioned unfortunate situations, when additional measurements can (and do) put the system at risk of not being approved.

Recommendations for future work. In the current chapter, we simply propose an idea. It is still necessary to develop this idea and to come up with the corresponding standards.

Chapter 3

Need for Optimal Distributed Measurement of Cumulative Quantities Explains the Ubiquity of Absolute and Relative Error Components

In this chapter, we deal with the second of the above-described challenges– related to the need for optimal organization of measurements. Specifically, we deal with the simplest case of this challenge, when the data processing algorithms consists of simply adding or averaging the measurement results. This case is practically important, since in many practical situations, we need to measure the value of a cumulative quantity, i.e., a quantity that is obtained by adding measurement results corresponding to different spatial locations. How can we select the measuring instruments so that the resulting cumulative quantity can be determined with known accuracy – and, to avoid unnecessary expenses, not more accurately than needed? It turns out that the only case where such an optimal arrangement is possible is when the required accuracy means selecting the upper bounds on absolute and relative error components. These results provide a possible explanation for the ubiquity of such two-component accuracy requirements.

3.1 Formulation of the Problem

Need for distributed measurements. In many practical situations, we are interested in estimating the value x of a cumulative quantity: e.g., we want to estimate the overall amount of oil in a given area, the overall amount of CO_2 emissions, etc.

How to perform distributed measurements. Measuring instruments usually measure quantities in a given location, i.e., they measure local values $x_1, \ldots, x_n$ that together form the desired value

$$x = x_1 + \ldots + x_n.$$

So, a natural way to produce an estimate $\widetilde{x}$ for the cumulative value x is:

- to place measuring instruments at several locations within a given area,
- to measure the values x_i of the desired quantity in these locations, and
- to add up the results $\widetilde{x}_1 + \ldots + \widetilde{x}_n$ of these measurement:

$$\widetilde{x} = \widetilde{x}_1 + \ldots + \widetilde{x}_n.$$

Need for optimal planning. Usually, we want to reach a certain estimation accuracy. To achieve this accuracy, we need to plan how accurate the deployed measurement instruments should be. Use of accurate measuring instruments is often very expensive, while budgets are usually limited. It is therefore desirable to come up with the deployment plan that would achieve the desired overall accuracy within the minimal cost. This implies, in particular, that the resulting estimate should not be more accurate than needed – this would mean that we could use less accurate (and thus, cheaper) measuring instruments.

What we do in this chapter. In this chapter, we provide a condition under which such optimal planning is possible – and the corresponding optimal planning algorithm. The resulting condition will explain why usually, measuring instruments are characterized by their absolute and relative accuracy.

3.2 Let Us Formulate the Problem in Precise Terms

How we can describe measurement accuracy. Measurements are never absolutely accurate, the measurement result $\widetilde{x}_i$ is, in general, different from the actual (unknown) value x_i of the corresponding quantity. In other words, the difference $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$ is, in general, different from 0. This difference is known as the *measurement error*.

For each measuring instrument, we should know how large the measurement error can be. In precise terms, we need to know an upper bound Δ on the absolute value $|\Delta x_i|$ of the measurement error. This upper bound should be provided by the manufacturer of the measuring instrument. Indeed, if no such upper is known, this means that whatever the reading of the measuring instrument, the actual value can be as far away from this reading as possible, so we get no information whatsoever about the actual value – in this case, this is not a measuring instrument, it is a wild guess.

Ideally, in addition to knowing that the measurement error Δx_i is somewhere in the interval $[-\Delta, \Delta]$, it is desirable to know how probable are different values from this interval, i.e., what is the probability distribution on the measurement error. Sometimes, we know this probability distribution, but in many practical situations, we don't know it, and the upper bound is all we know. So, in this section, we will consider this value as the measure of the instrument's accuracy.

This upper bound Δ may depend on the measured value. For example, if we are measuring current in the range from 1 mA to 1 A, then it is relatively easy to maintain accuracy of 0.1 mA when the actual current is 1 mA – this means measuring with one correct decimal digit. We can get values $0.813\ldots$, $0.825\ldots$, but since the measurement accuracy is 0.1, this means that these measurement results may correspond to the same actual value. In other words, whatever the measuring instrument shows, only one digit is meaningful and significant – all the other digits may be caused by measurement errors. On the other hand, to maintain the same accuracy of 0.1 mA when we measure currents close to 1 A would mean that we need to distinguish between values 0.94651 A = 946.51 mA and

0.94637 A = 946.37 mA, since the difference between these two values is larger than 0.1 mA. This would mean that we require that in the measurement result, we should have not one, but four significant digits – and this would require much more accurate measurements.

Because of this, we will explicitly take into account that the accuracy Δ depends on the measured value: $\Delta = \Delta(x)$. Usually, small changes in x lead to only small changes in the accuracy, so we can safely assume that the dependence $\Delta(x)$ is smooth.

What we want. We want to estimate the desired cumulative value x with some accuracy δ. In other words, we want to make sure that the difference between our estimate $\widetilde{x}$ and the actual value x does not exceed δ: $|\widetilde{x} - x| \le \delta$.

The cumulative value is estimated based on n measurement results. As we have mentioned, the accuracy that we can achieve in each measurement, in general, depends on the measured value: the larger the value of the measured quantity, the more difficult it is to maintain the corresponding accuracy. It is therefore reasonable to conclude that, whatever measuring instruments we use to measure each value x_i, it will be more difficult to estimate the larger cumulative value x with the same accuracy. Thus, it makes sense to require that the desired accuracy δ should also depend on the value that we want to estimate $\delta = \delta(x)$: the larger the value x, the larger the uncertainty $\delta(x)$ that we can achieve.

So, our problem takes the following form:

- we want to be able to estimate the cumulative value x with given accuracy $\delta(x)$ – i.e., we are given a function $\delta(x)$ and we want to estimate the cumulative value with this accuracy;
- we want to find the measuring instruments that would guarantee this estimation accuracy – and that would be optimal for this task, i.e., that would not provide better accuracy than needed.

Let us describe what we want in precise terms. To formulate this problem in precise terms, let us analyze what estimation accuracy we can achieve if we use, for each of n measurements, the measuring instrument characterized by the accuracy $\Delta(x)$.

Based on each measurement result $\widetilde{x}_i$, we can conclude that the actual value x_i of the corresponding quantity is located somewhere in the interval $[\widetilde{x}_i - \Delta(x_i), \widetilde{x}_i + \Delta(x_i)]$: the smallest possible value is $\widetilde{x}_i - \Delta(x_i)$, the largest possible value is $\widetilde{x}_i + \Delta(x_i)$.

When we add the measurement results, we get the estimate $\widetilde{x} = \widetilde{x}_1 + \ldots + \widetilde{x}_n$ for the desired quantity x. What are the possible values of this quantity? The sum $x = x_1 + \ldots + x_n$ attains its smallest value if all values x_i are the smallest, i.e., when

$$x = (\widetilde{x}_1 - \Delta(x_1)) + \ldots + (\widetilde{x}_n - \Delta(x_n)) = (\widetilde{x}_1 + \ldots + \widetilde{x}_n) - (\Delta(x_1) + \ldots + \Delta(x_n)),$$

i.e., when

$$x = \widetilde{x} - (\Delta(x_1) + \ldots + \Delta(x_n)).$$

Similarly, the sum $x = x_1 + \ldots + x_n$ attains its largest value if all values x_i are the largest, i.e., when

$$x = (\widetilde{x}_1 + \Delta(x_1)) + \ldots + (\widetilde{x}_n + \Delta(x_n)) = (\widetilde{x}_1 + \ldots + \widetilde{x}_n) + (\Delta(x_1) + \ldots + \Delta(x_n)),$$

i.e., when

$$x = \widetilde{x} + (\Delta(x_1) + \ldots + \Delta(x_n)).$$

Thus, all we can conclude about the value x is that this value belongs to the interval

$$[\widetilde{x} - (\Delta(x_1) + \ldots + \Delta(x_n)), \widetilde{x} + (\Delta(x_1) + \ldots + \Delta(x_n))].$$

This means that we get an estimate of x with the accuracy $\Delta(x_1) + \ldots + \Delta(x_n)$.

Our objective is to make sure that this is exactly the desired accuracy $\delta(x)$. In other words, we want to make sure that whenever $x = x_1 + \ldots + x_n$, we should have

$$\delta(x) = \Delta(x_1) + \ldots + \Delta(x_n).$$

Substituting $x = x_1 + \ldots + x_n$ into this formula, we get

$$\delta(x_1 + \ldots + x_n) = \Delta(x_1) + \ldots + \Delta(x_n). \tag{3.1}$$

We do not know a priori what will be the values x_i, so if we want to maintain the desired accuracy $\delta(x)$ – and make sure that we do not get more accuracy – we should make sure that the equality (3.1) be satisfied for all possible values $x_1, \ldots, x_n$.

In these terms, the problem takes the following form:

- For which functions $\delta(x)$ is it possible to have a function $\Delta(x)$ for which the equality (3.1) is satisfied? and
- For the functions $\delta(x)$ for which such function $\Delta(x)$ is possible, how can we find this function $\Delta(x)$ – that describes the corresponding measuring instrument?

This is the problem that we solve in this chapter.

3.3 When Is Optimal Distributive Measurement of Cumulative Quantities Possible?

Let us first analyze when the optimal distributive measurement of a cumulative quantity is possible, i.e., for which functions $\delta(x)$, there exists a function $\Delta(x)$ for which the equality (3.1) is always satisfied.

We have assumed that the function $\Delta(x)$ is smooth, i.e., differentiable. Thus, the sum $\delta(x)$ of such functions is differentiable too. Since both functions $\Delta(x)$ and $\delta(x)$ are differentiable, we can differentiate both sides of the equality (3.1) with respect to one of the variables – e.g., with respect to the variable x_1. The terms $\Delta(x_1), \ldots, \Delta(x_n)$ do not depend on x_1 at all, so their derivative with respect to x_1 is 0, and the resulting formula takes the form

$$\delta'(x_1 + \ldots + x_n) = \Delta'(x_1), \tag{3.2}$$

where, as usual, δ' and Δ' denote the derivatives of the corresponding functions.

The equality (3.2) holds for all possible values $x_2, \ldots, x_n$. For every real number x_0, we can take, e.g., $x_2 = x_0 - x_1$ and $x_3 = \ldots + x_n = 0$, then we will have $x_1 + \ldots + x_n = x_0$,

and the equality (3.2) takes the form

$$\delta'(x_0) = \Delta'(x_1).$$

The right-hand side does not depend on x_0, which means that the derivative $\delta'(x_0)$ is a constant not depending on x_0 either.

The only functions whose derivative is a constant are linear functions, so we conclude that the dependence $\delta(x)$ is linear:

$$\delta(x) = a + b \cdot x$$

for some constants a and b.

Interestingly, this fits well with the usual description of measurement error, as consisting of two components:

- the absolute error component a that does not depend on x at all, and
- the relative error component – according to which, the bound on the measurement error is a certain percentage of the actual value x, i.e., has the form $b \cdot x$ for some constant b (e.g., for 10% accuracy, $b = 0.1$).

Thus, our result explains this usual description.

3.4 What Measuring Instrument Should We Select to Get the Optimal Distributive Measurement of Cumulative Quantity?

Now that we know for what desired accuracy $\delta(x)$, we can have the optimal distributive measurement of a cumulative quantity, the natural next question is: given one of such functions $\delta(x)$, what measuring instrument – i.e., what function $\Delta(x)$ – should we select for this optimal measurement?

To answer this question, we can take $x_1 = \ldots = x_n$. In this case, $\Delta(x_1) = \ldots = \Delta(x_n)$, so the equality (3.2) takes the form

$$\delta(n \cdot x_1) = n \cdot \Delta(x_1). \tag{3.3}$$

We know that $\delta(x) = a + b \cdot x$, so the formula (3.3) takes the form

$$a + b \cdot n \cdot x_1 = n \cdot \Delta(x_1).$$

If we divide both sides of this equality by x_1, and rename x_1 into x, we get the desired expression for $\Delta(x)$:

$$\Delta(x) = \frac{a}{n} + b \cdot x.$$

In other words:

- the bound on the relative error component of each measuring instrument should be the same as the desired relative accuracy of the cumulative quantity, and
- the bound on the absolute error component should be n times smaller than the desired bound on the absolute accuracy of the cumulative quantity.

3.5 Conclusions and Recommendations for Future Work

Conclusions. In this chapter, we deal with the second of the four metrological challenges listed in Chapter 1. This challenge is related to the fact that in the past, when there were few affordable measuring instruments and we could only afford a few measurements, there were not that many options. In such cases, planning measurements simply meant selecting one of these options. So, we could plan the measurements "by hand". Nowadays, with a potential to perform a large number of measurements and the availability of many different measuring instruments, the number of possible measurement options becomes so large that we need to develop methods for optimal planning. There exist techniques for such

planning, but these techniques are mostly based on limited number of measurements. For situations when we have a large number of measurements, to the best of our knowledge, no practical general methods are known – even for the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results.

In the current chapter, we provide a theoretical analysis of the problem and find a new explicit formulas for the optimal measurement design. As an interesting side effect of this theoretical analysis, we come up with an explanation of why measurement accuracy is usually described by listing absolute and relative error components. To the best of our knowledge, ours is the first theoretical explanation for this widely used practice.

Recommendations for future work. In this chapter, we only deal with the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results. It is necessary to extend our analysis to more complex data processing algorithms.

Chapter 4

Graph Approach to Uncertainty Quantification

In the previous chapters, we outlined four main challenges of practical computer-enhanced measurements (in Chapter 1), and suggested – in Chapters 2 and 3 – how to deal with the first two challenges. In this chapter, we deal with the third challenge: that for different pairs of measurements, we may have different information about the dependence between the corresponding measurement errors. Specifically, we develop techniques for processing measurement results in situations which are slightly different from the above-described well-studied ones.

Traditional analysis of uncertainty of the result of data processing assumes that all measurement errors are independent. In reality, there may be common factor affecting these errors, so these errors may be dependent. In such cases, the independence assumption may lead to underestimation of uncertainty. In such cases, a guaranteed way to be on the safe side is to make no assumption about independence at all. In practice, however, we may have information that a few pairs of measurement errors are indeed independent – while we still have no information about all other pairs. Alternatively, we may suspect that for a few pairs of measurement errors, there may be correlation – but for all other pairs, measurement errors are independent. In both cases, unusual pairs can be naturally represented as edges of a graph. In this chapter, we show how to estimate the uncertainty of the result of data processing when this graph is small.

4.1 Introduction

What is the problem and what we do about it: a brief description. Estimating uncertainty of the result of data processing is important in many practical applications. Corresponding formulas are well known for two extreme cases:

- when all measurement errors are independent, and
- when we have no information about the dependence.

These cases are indeed ubiquitous, but often, the actual cases are somewhat different; e.g.:

- most pairs of inputs are known to be independent, but
- there are a few pairs for which we are not sure.

Alternatively, for almost all pairs, we may have no information about the dependence, but for a few pairs of inputs, we know that the corresponding measurement errors are independent. Such unusual pairs can be naturally represented as edges of a graph. It is desirable to analyze how the presence of this graph changes the corresponding estimates.

In this chapter, we start answering this question for all graphs of sizes 2, 3, and 4. We hope that our results will be extended to larger-size graphs.

Structure of the chapter. In Section 2, we provide a detailed description of the general problem, and describe how uncertainty is estimated in the above-described two extreme cases. In the following sections, we present our results about situations in which the deviation from one of these extreme cases is described by a small-size graph.

4.2 Detailed Formulation of the Problem

Need for data processing. One of the main objectives of science is to describe the current state of the world and to predict future events based on what we know about the

current and past states. In general, the state of a system is characterized by the values of corresponding quantities.

Some quantities we can measure directly – e.g., we can directly measure the temperature in the room or the distance between two campus buildings. However, some quantities cannot (yet) be measured directly: we cannot directly measure the temperature inside a star or a distance to this star. Since we cannot measure such quantities directly, the only way we can estimate the values of these quantities is by measuring them indirectly: i.e., by measuring related quantities $x_1, \ldots, x_n$ that are related by y by a known dependence $y = f(x_1, \ldots, x_n)$. Once we know such related quantities, we can measure their values, and use the measurement results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ to compute the estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for y. Computing this estimate is an important case of *data processing*.

Data processing is also needed for *predictions*. For example, we may want to predict the future location of a near-Earth asteroid or tomorrow's weather. The future state can be described if we describe the future values of all the quantities characterizing this state. For example, tomorrow's weather can be characterized by temperature, wind speed, etc. To be able to make this prediction, for each of the quantities describing the future state, we need to know the relation $y = f(x_1, \ldots, x_n)$ between the future value y of this quantity and the current and past values x_i of related quantities. Once we know this relation, then we can use it to transform the measured values $\widetilde{x}_i$ of the quantities x_i into the estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for the desired quantity y. Computing $\widetilde{y}$ based on the measured values $\widetilde{x}_i$ is another important case of data processing.

Need for uncertainty quantification. Measurement results $\widetilde{x}$ are, in general, somewhat different from the actual (unknown) value x of the corresponding quantity; see, e.g., [20]. In other words, the difference $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ is usually non-zero. This difference is known as the *measurement error*.

Since the inputs $\widetilde{x}_i$ to the algorithm f are, in general, different from the actual values x_i, the resulting estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ is, in general, different from the actual value $y = f(x_1, \ldots, x_n)$ that we would have gotten if we knew the exact values x_i. In practical

situations, it is important to know how big this difference can be. For example, suppose we predict that the asteroid will pass at a distance of 150,000 km from the Earth; then:

- if the accuracy of this estimate is $\pm 200,000$ km, then this asteroid may hit the Earth, while
- if the the accuracy is $\pm 20,000$ km, this particular asteroid is harmless.

Estimating the accuracy of our estimates is an important case of *uncertainty quantification*.

What we know about measurement errors. In similar situations, with the exact same value of the measured quantity, the same measuring instrument can produce different results. This is well known to anyone who has ever repeatedly measured the same quantity: the results are always somewhat different, whether it is a current or body temperature or blood pressure. In this sense, measurement errors are *random*. Each random variable has an average (mean) value, and its actual values deviate from this mean.

Measuring instruments are usually calibrated: the measurement results provided by this instrument are compared with measurement results provided by a much more accurate ("standard") measuring instrument. If the mean difference is non-zero – i.e., in statistical terms, if the measuring instrument has a *bias* – then we can simply subtract this bias from all the measurement results and thus, make the mean error equal to 0. For example, if a person knows that his/her watch is 5 minutes ahead, this person can always subtract 5 minutes from the watch's reading and get the correct time. So, we can safely assume that the mean value $E[\Delta x]$ of each measurement error Δx is 0: $E[\Delta x] = 0$.

The deviations from the mean value are usually described by the mean squared deviation – which is known as the *standard deviation* σ. Instead of the standard deviation σ, it is sometimes convenient to use its square $V \stackrel{\text{def}}{=} \sigma^2$ which is called the *variance*. In precise terms, the variance is the mean value of the square of the difference between the random variable and its mean value: $V[X] = E\left[(X - E[X])^2\right]$. For measurement error, the mean is $E[\Delta x] = 0$, so we get a simplified formula $V[\Delta x] = E[(\Delta x)^2]$.

For each measuring instrument, the standard deviation is also determined during the calibration. So, we can assume that for each measuring instrument:

- we know that the mean value of its measurement error is 0, and
- we know the standard deviation of the measurement error.

In many cases, distributions are normal. In most practical cases, there are many factors that contribute to the measurement error. For example, if we measure voltage, the measuring instrument is affected not only by the current that we measure but also by the currents of multiple devices present in the room, including the computer used to process the data, the lamps in the ceiling, etc. Each of these factors may be relatively small, but there are many of them, and thus, the resulting measurement error is much larger than each of them.

It is known – see, e.g., [23] – that the probability distribution of the joint effect of a large number of small random factors is close to Gaussian (normal). Thus, in such cases, we can safely assume that the measurement errors are normally distributed.

Possibility of linearization. In general, the estimation error is equal to $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$. Here, $\widetilde{y} = f\,(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ and $y = f\,(x_1, \ldots, x_n)$, so

$$\Delta y = f\,(\widetilde{x}_1, \ldots, \widetilde{x}_n) - f\,(x_1, \ldots, x_n)\,,$$

By definition of the measurement error Δx_i as the difference $\Delta x_i = \widetilde{x}_i - x_i$, we have $x_i = \widetilde{x}_i - \Delta x_i$. Thus, the above expression for the approximation error takes the form

$$\Delta y = f\,(\widetilde{x}_1, \ldots, \widetilde{x}_n) - f\,(\widetilde{x}_1 - \Delta x_1, \ldots, \widetilde{x}_n - \Delta x_n)\,. \tag{4.1}$$

Measurement errors are usually relatively small, so that terms quadratic in these errors can be safely ignored. For example, if the measurement error is 10%, its square is 1%, which is much smaller. Thus, we can expand the right-hand side of the equality (4.1) in Taylor series and keep only linear terms in this expansion. As a result, we conclude that

$$\Delta y = \sum_{i=1}^{n} c_i \cdot \Delta x_i, \tag{4.2}$$

where we denoted

$$c_i \stackrel{\text{def}}{=} \frac{\partial f}{\partial x_i}_{|x_1=\widetilde{x}_1,\ldots,x_n=\widetilde{x}_n}.$$

In other words, the desired estimation error Δy is a linear combination of measurement errors Δx_i.

Case when all measurement errors are independent. It is known that the variance of the sum of the several random variables is equal to the sum of their variances. It is also known that if we multiply a random variable by a constant, then its standard deviation is multiplied by the absolute value of this constant. So, if we denote the standard deviation of the i-th measuring instrument by σ_i, then the standard deviation of the product $c_i \cdot \Delta x_i$ is equal to $|c_i| \cdot \sigma_i$ and thus, its variance is equal to $(|c_i| \cdot \sigma_i)^2 = c_i^2 \cdot \sigma_i^2$. Thus, the variance of the sum Δy is equal to the sum of these variances:

$$\sigma^2 = \sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2, \tag{4.3}$$

and thus, the standard deviation is equal to

$$\sigma = \sqrt{\sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2}. \tag{4.4}$$

Towards the general case: a known geometric interpretation of random variables. We have n random variables $v_i \stackrel{\text{def}}{=} c_i \cdot \Delta x_i$. For each variable, we know its standard deviation $|c_i| \cdot \sigma_i$, and we are interested in estimating the standard deviation of the sum $\Delta y = v_1 + \ldots + v_n$ of these variable. It is known (see, e.g., [23]) that we can interpret each variable – and, correspondingly, each linear combination of the variables – as vectors $\vec{a}$, $\vec{b}$ in an n-dimensional space, so that the length $\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}}$ of each vector (where $\vec{a} \cdot \vec{b}$ means dot (scalar) product) is equal to the standard deviation of the corresponding variable. In these terms, independence means that the two vectors are orthogonal. Indeed:

- In statistical terms, independence implies that the variance of the sum if equal to the sum of the variances.

- For the sum $\vec{a}+\vec{b}$ of two vectors, the square of the length has the form

$$\left\|\vec{a}+\vec{b}\right\|^2 = \left(\vec{a}+\vec{b}\right)\cdot\left(\vec{a}+\vec{b}\right) = \vec{a}\cdot\vec{a}+\vec{b}\cdot\vec{b}+2\vec{a}\cdot\vec{b}.$$

  Here, $\vec{a}\cdot\vec{a} = \|\vec{a}\|^2$, $\vec{b}\cdot\vec{b} = \left\|\vec{b}\right\|^2$, and $\vec{a}\cdot\vec{b} = \|\vec{a}\|\cdot\left\|\vec{b}\right\|\cdot\cos(\theta)$, where θ is the angle between the two vectors, so the above expression takes the form

$$\left\|\vec{a}+\vec{b}\right\|^2 = \|\vec{a}\|^2 + \left\|\vec{b}\right\|^2 + 2\|\vec{a}\|\cdot\left\|\vec{b}\right\|\cdot\cos(\theta).$$

  So, the variance $\left\|\vec{a}+\vec{b}\right\|^2$ of the sum is equal to the sum $\|\vec{a}\|^2+\left\|\vec{b}\right\|^2$ of the variances if and only if $\cos(\theta)=0$, i.e., if only if the angle is 90°, and the vectors are orthogonal.

In the independent case, n vectors $\vec{v}_i$ corresponding to individual measurements are orthogonal to each other, so, similarly to the above argument, one can show that the length of their sum is equal to the square root of the sum of the squares of their lengths:

$$\|\vec{v}_1+\ldots+\vec{v}_n\|^2 = \|\vec{v}_1\|^2+\ldots+\|\vec{v}_n\|^2.$$

Let us use this geometric interpretation to estimate the uncertainty in situations when we know nothing about correlation between different measurement errors.

What if we have no information about correlations: analysis of the problem and the resulting formula. In this case, we still have n vectors $\vec{v}_1,\ldots,\vec{v}_n$ of given lengths $\|\vec{v}_i\| = |c_i|\cdot\sigma_i$. The main difference from the independent case is that these vectors are not necessarily orthogonal, we can have different angles between them. In this case, in contrast to the independent case, the length of the sum is not uniquely determined. For example:

- if two vectors of equal length are parallel, the length of their sum is double the length of each vector, but
- if they are anti-parallel $\vec{b}=-\vec{a}$, then their sum has length 0.

In such cases, it is reasonable to find the worst possible standard deviation, i.e., the largest possible length.

One can easily check that the sum of several vectors of given length is the largest when all these vectors are parallel and oriented in the same direction. In this case, the length of the sum is simply equal to the sum of the lengths, so we get

$$\sigma = \sum_{i=1}^{n} |c_i| \cdot \sigma_i, \tag{4.5}$$

and

$$\sigma^2 = \left(\sum_{i=1}^{n} |c_i| \cdot \sigma_i \right)^2. \tag{4.6}$$

Precise mathematical formulation of this result. The above result can be presented in the following precise form.

Definition 4.1.

- *Let $s = (\sigma_1, \ldots, \sigma_n)$ be a tuple of non-negative real numbers.*
- *Let D denote the class of all possible multi-D distributions $(\Delta x_1, \ldots, \Delta x_n)$ for which, for each i, we have $E[\Delta x_i] = 0$ and $\sigma[\Delta x_i] = \sigma_i$.*
- *Let $\mathcal{S}$ be a subset of the set D; we will denote it, as usual, by $\mathcal{S} \subseteq D$.*
- *Let $c = (c_1, \ldots, c_n)$ be a tuple of real numbers.*
- *For each distribution from D, let Δy denote $\Delta y \stackrel{\text{def}}{=} c_1 \cdot \Delta x_1 + \ldots + c_n \cdot \Delta x_n$.*

Then, by $\sigma(c, s, \mathcal{S})$ we denote the largest possible value of the standard deviation $\sigma_\rho[\Delta y]$ over all distributions from the set $\mathcal{S}$:

$$\sigma(s, \mathcal{S}, c) \stackrel{\text{def}}{=} \max_{\rho \in \mathcal{S}} \sigma_\rho[\Delta y].$$

In these terms, the above result takes the following form:

Proposition 4.1. *For the set $\mathcal{S} = D$ of all possible distributions, we have*

$$\sigma\,(s, D, c) = \sum_{i=1}^{n} |c_i| \cdot \sigma_i.$$

Similarly, the previous result – about independent case – takes the following form.

Definition 4.2. *By $I \subset D$, we denote the class of all distributions for which, for all i and j, the variables Δx_i and Δx_j are independent. We will call I* independent set.

Proposition 4.2. *For the independent set I, we have*

$$\sigma\,(s, I, c) = \sqrt{\sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2}.$$

Comment. Interestingly, the formula (4.5) is similar to what we get in the linearized version of the interval case (see, e.g., [10, 13, 16, 19, 20]), i.e., the case when we only know the upper bound Δx_i on the absolute value of the measurement error. In other words, this means that:

- the measurement error is located somewhere in the interval $[-\Delta x_i, \Delta x_i]$, and
- we have no information about the probability of different values from this interval.

In this case, the largest possible value of the estimation error

$$\Delta y = c_1 \cdot \Delta x_1 + \ldots + c_n \cdot \Delta x_n$$

is equal to $|c_1| \cdot \Delta_1 + \ldots + |c_n| \cdot \Delta_n$. This is indeed the same expression as our formula (4.5).

4.3 What If a Few Pairs of Measurement Errors Are Not Necessarily Independent

4.3.1 Description of the Situation

Descriptions of the situation. In the previous text, we considered two extreme cases:

- when we know that all measurement errors are independent, and
- when we have no information about their dependence.

Such cases are indeed frequent, but sometimes, situations are similar but not exactly the same. For example, we can have the case of "almost independence", when for most pairs, we know that they are independent, but for a few pairs, we do not have this information. This is the situation that we describe in this section.

Comment. The opposite situations, when we only have independence information about a few pairs, is described in the next section.

Graph representation of such situations. To describe such situations, we need to know for which pairs of measurement errors, we do not have information about independence. A natural way to represent such information is by an undirected graph in which:

- measurement errors are vertices and
- an edge connects pairs for which we do not have information about independence.

We only need to know which vertices are connected. So, it makes sense to include, in the description of the graph, only vertices that are connected by some edge, i.e., only measurement errors that may not be independent with respect to others. In this case, we arrive at the following definition.

Definition 4.3.

- *Let $G = (V, E)$ be an undirected graph with the set of vertices $V \subseteq \{1, \ldots, n\}$ for which every vertex is connected to some other vertex. Here, E is a subset $E \subseteq V \times V$ for which:*
  - *for each $a \in V$, we have $(a, a) \notin E$,*
  - *for each a and b, $(a, b) \in E$ if and only if $(b, a) \in E$, and*
  - *for each $a \in V$, we have $(a, b) \in E$ for some $b \in V$.*

- *By $I_G \subseteq D$, we mean the class of all distributions for which for all pairs $(i, j) \notin E$ the variables Δx_i and Δx_j are independent.*

Discussion. Our objective is to find the value $\sigma(s, I_G, c)$ for different graphs G. In this chapter, we only consider the simplest graphs: all graphs with 2, 3, or 4 vertices. We hope that this work will be extended to larger-size graphs.

4.3.2 General Results

Let us first present some general results. For this purpose, let us introduce the following notations. For any set $S \subseteq \{1, \ldots, n\}$, by a restriction c_S we mean sub-tuples consisting only of elements c_i for which $i \in S$. For example, for $c = (c_1, c_2, c_3, c_4)$ and $S = \{1, 3\}$, we have $c_S = (c_1, c_3)$. Similarly, we can define the restriction s_S. It is then relatively easy to show that the following result holds:

Proposition 4.3. *For each graph $G = (V, E)$, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + (\sigma(s_V, I_G, c_V))^2.$$

Comments.

- In other words, it is sufficient to only consider measurement errors from the exception set V – which are not necessarily independent, then all other measurement errors can be treated the same way as in the case when all measurement errors are independent.
- For reader's conveniences, all the proofs are placed in a special Proofs section.

Another easy-to-analyze important case is when the graph G is disconnected, i.e., consists of several connected components.

Proposition 4.4. *When the graph $G = (V, E)$ consists of several connected components $G_1 = (V_1, E_1), \ldots, G_k = (V_k, E_k)$ with for which $V = V_1 \cup \ldots \cup V_k$, then*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \sum_{j=1}^{k} \left(\sigma\left(s_{V_j}, \mathcal{S}_{G_j}, c_{V_j}\right)\right)^2.$$

Comment. In view of this result, it is sufficient to estimate the value $\sigma(s, I_G, c)$ for connected graphs. We consider connected graphs with 2, 3, or 4 vertices.

4.3.3 Connected Graph of Size 2

There is only one connected graph of size 2: two vertices i_1 and i_2 connected by an edge, so that $V = \{i_1, i_2\}$ and $E = \{(i_1, i_2), (i_2, i_1)\}$.

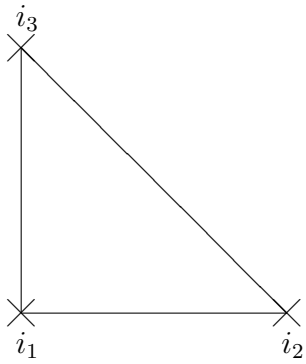
Proposition 4.5. *When the graph $G = (V, E)$ consists of two vertices i_1 and i_2 connected by an edge, then*

$$\sigma^2(s, I_G, c) = \sum_{i \neq i_1, i_2} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + |c_{i_2}| \cdot \sigma_{i_2}\right)^2.$$

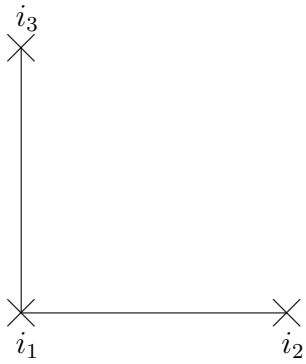
4.3.4 Connected Graphs of Size 3

In a connected graph of size 3, two vertices are connected, and the third vertex is:

- either connected to both of them – in this case we have a connected 3-element graph,



- or to only one of them.



So, modulo isomorphism, there are two different connected graphs of size 3. For these graphs, we get the following results:

Proposition 4.6. *When $G = (V, E)$ is a complete 3-element graph with $V = \{i_1, i_2, i_3\}$, then*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + |c_{i_2}| \cdot \sigma_{i_2} + |c_{i_3}| \cdot \sigma_{i_3}\right)^2.$$

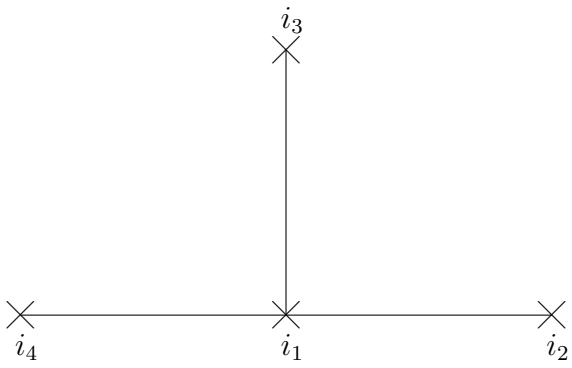
Proposition 4.7. *For a 3-element graph with $V = \{i_1, i_2, i_3\}$ in which i_1 is connected to i_2 and i_3 but i_2 and i_3 are not connected, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + \sqrt{c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2}\right)^2.$$

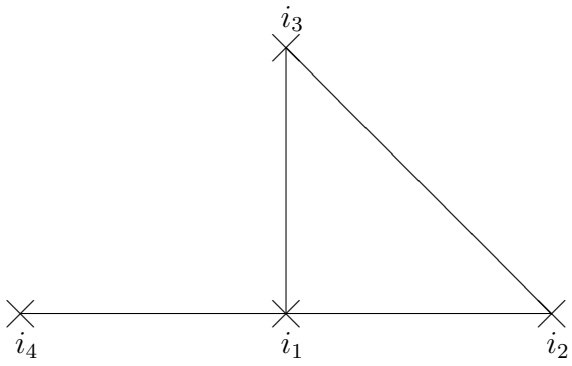
4.3.5 Connected Graphs of Size 4

Let us first consider graphs of size 4 for which there is a vertex (we will denote it i_1) connected with all three other vertices. In this case, there are four possible options:

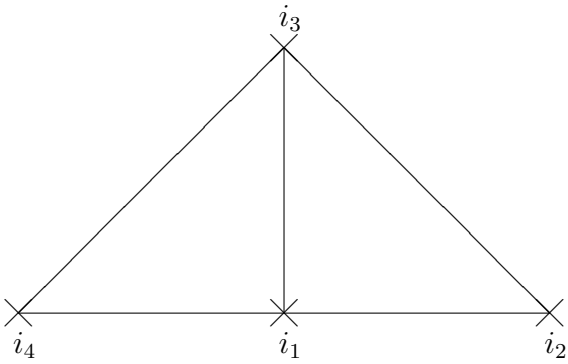
- when connections between i_1 and all three other vertices are the only connections:



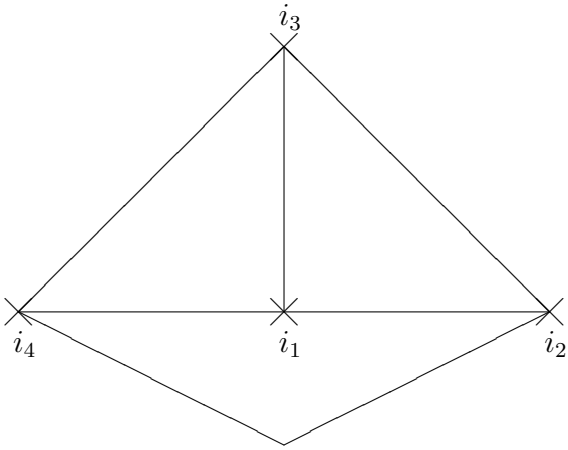
- when, in addition to edges connecting i_1 to three other vertices, there is also one connection between two of these other vertices:



- when, in addition to edges connecting i_1 to three other vertices, there are two connection between these other vertices:



- and when we have a complete 4-element graph:

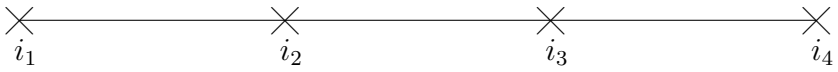


Finally, let us consider graphs in which each vertex is connected to no more than two others. If each vertex is connected to only one vertex, then a vertex i_1 is connected to some vertex i_2, and there is no space for each of them to have any other connection – so the 4-element graph containing vertices i_1 and i_2 cannot be connected. Thus, there should be at least one vertex connected to two others.

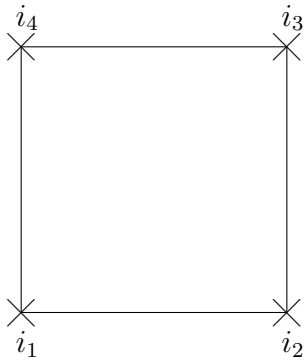
Let us denote one such vertex by i_2, and the two vertices to which i_2 is connected by i_1 and i_3. Since the graph is connected, the fourth vertex i_4 must be connected to one of the previous three vertices. The vertex i_4 cannot be connected to i_2 – because then i_2 should be connected to three other vertices, and we consider the case when each vertex is connected

to no more than two others. Thus, i_4 is connected to i_1 and/or i_3. If it is connected to i_1, then we can swap the names of vertices i_1 and i_3, and get the same configuration as when i_4 is connected to i_3. If i_4 is connected to both i_1 and i_3, then the resulting graph is uniquely determined. Thus, under the assumption that each vertex is connected to no more than two others, we have two possible graphs:

- a "linear" graph:



- and a "square graph":



For all these graphs, we have the following results:

Proposition 4.8. *For a 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, in which i_1 is connected to i_2, i_3, and i_4, but i_2, i_3, and i_4 are not connected to each other, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + \sqrt{c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}\right)^2.$$

Proposition 4.9. *For a 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, in which i_1, i_2, and i_3 form a complete graph, and i_4 is connected only to i_1, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + \sqrt{(|c_{i_2}| \cdot \sigma_{i_2} + |c_{i_3}| \cdot \sigma_{i_3})^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}\right)^2.$$

Proposition 4.10. *For a 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, in which i_1, i_2, and i_3 form a complete graph, and i_4 is corrected to i_1 and i_3, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + |c_{i_3}| \cdot \sigma_{i_3} + \sqrt{c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}\right)^2.$$

Proposition 4.11. *For a complete 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(|c_{i_1}| \cdot \sigma_{i_1} + |c_{i_2}| \cdot \sigma_{i_2} + |c_{i_3}| \cdot \sigma_{i_3} + |c_{i_4}| \cdot \sigma_{i_4}\right)^2.$$

Proposition 4.12. *For a linear 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, the value $\sigma(s, I_G, c)$ has the following form:*

- *if $|c_{i_2}| \cdot \sigma_{i_2} \cdot |c_{i_3}| \cdot \sigma_{i_3} > |c_{i_1}| \cdot \sigma_{i_1} \cdot |c_{i_4}| \cdot \sigma_{i_4}$, then*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \left(\sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_2}^2 \cdot \sigma_{i_2}^2} + \sqrt{c_{i_3}^2 \cdot \sigma_{i_3}^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}\right)^2;$$

- *otherwise, we get*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 + \max(\sigma_2^2, \sigma_3^2, \sigma_0^2), \text{ where}$$

$$\sigma_2^2 = v_{i_3}^2 + \left(\sqrt{v_{i_1}^2 + v_{i_2}^2} + v_{i_4}\right)^2,$$

$$\sigma_3^2 = v_{i_2}^2 + \left(v_{i_1} + \sqrt{v_{i_3}^2 + v_{i_4}^2}\right)^2, \text{ and}$$

$$\sigma_0^2 = (v_{i_1} + v_{i_3})^2 + (v_{i_2} + v_{i_4})^2.$$

Proposition 4.13. *For a square 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, we have:*

$$\sigma^2(s, I_G, c) = \sum_{i \notin V} c_i^2 \cdot \sigma_i^2 +$$

$$\left(\sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2} + \sqrt{c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}\right)^2.$$

4.4 What If Only a Few Pairs of Measurement Errors Are Known to Be Independent

4.4.1 Description of the Situation

Graph representation of such situations. To describe such situations, we need to know for which pairs of measurement errors, we have information about independence. A natural way to represent such information is by an undirected graph in which:

- measurement errors are vertices and
- an edge connects pairs for which we have information about independence.

For simplicity, we can only consider vertices that are connected by some edge, i.e., only measurement errors that are known to be independent with respect to others. In this case, we arrive at the following definition.

Definition 4.4.

- *Let $G = (V, E)$ be an undirected graph with the set of vertices $V \subseteq \{1, \ldots, n\}$. Here, E is a subset $E \subseteq V \times V$ for which:*
  - *for each $a \in V$, we have $(a, a) \notin E$, and*
  - *for each a and b, $(a, b) \in E$ if and only if $(b, a) \in E$.*
- *By $D_G \subseteq D$, we mean the class of all distributions for which for all pairs $(i, j) \in E$ the variables Δx_i and Δx_j are independent.*

Discussion. Our objective is to find the value $\sigma(s, D_G, c)$ for different graphs G. In this chapter, we only consider the simplest graphs: all graphs with 2, 3, or 4 vertices. We hope that this work will be extended to larger-size graphs.

4.4.2 General Results

Let us first present some general results. It is then relatively easy to show that the following result holds:

Proposition 4.14. *For each graph $G = (V, E)$, we have:*

$$\sigma\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sigma\left(s_V, D_G, c_V\right).$$

Comment. In other words, it is sufficient to only consider measurement errors from the exception set V – which are not necessarily independent; then all other measurement errors can be treated the same way as in the case when we have no information about dependence.

Another easy-to-analyze important case is when the graph G is disconnected, consisting of several connected components.

Proposition 4.15. *When the graph $G = (V, E)$ consists several connected components $G_1 = (V_1, E_1), \ldots, G_k = (V_k, E_k)$ with for which $V = V_1 \cup \ldots \cup V_k$, then*

$$\sigma\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sum_{j=1}^{k} \sigma\left(s_{V_j}, \mathcal{S}_{G_j}, c_{V_j}\right).$$

Comment. In view of this result, it is sufficient to estimate the value $\sigma\left(s, D_G, c\right)$ for connected graphs. In this chapter, we consider all connected graphs with 2, 3, or 4 vertices.

4.4.3 Connected Graph of Size 2

Proposition 4.16. *When the graph $G = (V, E)$ consists of two vertices i_1 and i_2 connected by an edge, then*

$$\sigma^2\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_2}^2 \cdot \sigma_{i_2}^2}.$$

4.4.4 Connected Graphs of Size 3

Proposition 4.17. *When $G = (V, E)$ is a complete 3-element graph with $V = \{i_1, i_2, i_3\}$, then*

$$\sigma^2(s, D_G, c) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2}.$$

Proposition 4.18. *For a 3-element graph with $V = \{i_1, i_2, i_3\}$, in which i_1 is connected to i_2 and i_3 but i_2 and i_3 are not connected, we have:*

$$\sigma^2(s, D_G, c) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + \left(|c_{i_2}| \cdot \sigma_{i_2} + |c_{i_3}| \cdot \sigma_{i_3}\right)^2}.$$

4.4.5 Connected Graphs of Size 4

Proposition 4.19. *For a 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, in which i_1 is connected to i_2, i_3, and i_4, but i_2, i_3, and i_4 are not connected to each other, we have:*

$$\sigma(s, D_G, c) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + \left(|c_{i_2}| \cdot \sigma_{i_2} + |c_{i_3}| \cdot \sigma_{i_3} + |c_{i_4}| \cdot \sigma_{i_4}\right)^2}.$$

Proposition 4.20. *For a 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, in which i_1, i_2, and i_3 form a complete graph, and i_4 is connected only to i_1, we have:*

$$\sigma(s, D_G, c) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + \left(\sqrt{c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2} + |c_{i_4}| \cdot \sigma_{i_4}\right)^2}.$$

Proposition 4.21. *For a 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, in which i_1, i_2, and i_3 form a complete graph, and i_4 is corrected to i_1 and i_3, we have:*

$$\sigma(s, D_G, c) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2 + \left(|c_{i_2}| \cdot \sigma_{i_2} + |c_{i_4}| \cdot \sigma_{i_4}\right)^2}.$$

Proposition 4.22. *For a complete 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, , we have:*

$$\sigma\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_2}^2 \cdot \sigma_{i_2}^2 + c_{i_3}^2 \cdot \sigma_{i_3}^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}.$$

Proposition 4.23. *For a linear 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$:*

- *if $|c_{i_1}| \cdot \sigma_{i_1} \cdot |c_{i_4}| \cdot \sigma_{i_4} > |c_{i_2}| \cdot \sigma_{i_2} \cdot |c_{i_3}| \cdot \sigma_{i_3}$, then*

$$\sigma\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_2}^2 \cdot \sigma_{i_2}^2} + \sqrt{c_{i_3}^2 \cdot \sigma_{i_3}^2 + c_{i_4}^2 \cdot \sigma_{i_4}^2}.$$

- *otherwise, we get*

$$\sigma\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i + \sqrt{\max(\sigma_2^2, \sigma_3^2, \sigma_0^2)}, \text{ where}$$

$$\sigma_2^2 = v_{i_3}^2 + \left(\sqrt{v_{i_1}^2 + v_{i_2}^2} + v_{i_4}\right)^2,$$

$$\sigma_3^2 = v_{i_2}^2 + \left(v_{i_1} + \sqrt{v_{i_3}^2 + v_{i_4}^2}\right)^2, \text{ and}$$

$$\sigma_0^2 = \left(v_{i_1} + v_{i_3}\right)^2 + \left(v_{i_2} + v_{i_4}\right)^2.$$

Proposition 4.24. *For a square 4-element graph with $V = \{i_1, i_2, i_3, i_4\}$, we have:*

$$\sigma\left(s, D_G, c\right) = \sum_{i \notin V} |c_i| \cdot \sigma_i +$$

$$\sqrt{\left(|c_{i_1}| \cdot \sigma_{i_1} + |c_{i_3}| \cdot \sigma_{i_3}\right)^2 + \left(|c_{i_2}| \cdot \sigma_{i_2} + |c_{i_4}| \cdot \sigma_{i_4}\right)^2}.$$

4.5 Proofs

Proof of Proposition 4.3. The proof of this proposition naturally follows from the geometric interpretation, in which we associate a vector to each random variable, and we

are looking for a configuration in which the sum of these vectors has the largest length. Here, the sum $\vec{v}$ of all the corresponding vectors $\vec{v} = \vec{v}_1 + \ldots + \vec{v}_n$ can be represented as

$$\sum_{i \notin V} \vec{v}_i + \vec{a},$$

where

$$\vec{a} \stackrel{\text{def}}{=} \sum_{j \in V} \vec{v}_j.$$

Vectors $\vec{v}_i$ corresponding to "normal" errors ($i \notin V$) are orthogonal (since the corresponding measurement errors are independent), and since each of them is orthogonal to each of the "abnormal" vectors $\vec{v}_j$, it is also orthogonal to the sum $\vec{a}$ of these abnormal vectors. Thus, the square of the length of the sum $\vec{v}$ is equal to the sum of the squares of the lengths of the "normal" vectors $\vec{v}_i$ and of the vector $\vec{a}$:

$$\|\vec{v}\|^2 = \sum_{i \notin V} \|\vec{v}_i\|^2 + \|\vec{a}\|^2.$$

The values $\|\vec{v}_i\|^2$ are given: they are equal to $c_i^2 \cdot \sigma_i^2$ Thus, the largest possible value of $\|\vec{v}\|$ is attained when the length $\|\vec{a}\|$ is the largest. This largest length is what in Definitions 1 and 3 we denoted by $\sigma\,(s_V, I_G, c_V)$. Thus, we get the desired formula.

The proposition is proven.

Proof of Proposition 4.5. For this graph, the value $\sigma\,(s, I_G, c)$ follows from Proposition 4.1 – it is equal to $|c_{i_1}| \cdot \sigma_{i_1} + |c_{i_2}| \cdot \sigma_{i_2}$. Thus, by Proposition 4.3, we get the desired result.

Proof of Propositions 6 is similar to the proof of Proposition 4.5.

Proof of Proposition 4.7. Since the vertices i_2 and i_3 are not connected, this means that the measurement errors corresponding to these vertices are independent, so the length of $\vec{v}_{i_2} + \vec{v}_{i_3}$ is equal to $\sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2}$.

The vertex i_1 is connected to both i_2 and i_3, which means that we know nothing about the dependence between the corresponding measurement errors. Thus, as we have described earlier, the largest possible length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} = \vec{v}_{i_1} + (\vec{v}_{i_2} + \vec{v}_{i_3})$$

can be obtained by adding the lengths of $\vec{v}_{i_1}$ and of $\vec{v}_{i_2} + \vec{v}_{i_3}$:

$$\|\vec{v}_{i_3}\| + \sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2}.$$

The desired result now follows from Proposition 4.3.

Proof of Proposition 4.8 is similar to the proof of Proposition 4.7.

Proof of Proposition 4.9. We have no restriction on vectors $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$, so the largest possible length of their sum $\vec{v}_{i_2} + \vec{v}_{i_3}$ is the sum of their lengths: $\|\vec{v}_{i_2}\| + \|\vec{v}_{i_3}\|$. There is no edge between i_4 and the group of vertices (i_2, i_3), this means that the measurement errors corresponding to i_4 are independent from the measurement errors Δx_{i_2} and Δx_{i_3}. Thus, the vector $\vec{v}_{i_4}$ is orthogonal to vectors $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$ and is, thus, orthogonal to their sum $\vec{v}_{i_2} + \vec{v}_{i_3}$. So, the maximum length of the sum

$$\vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = (\vec{v}_{i_2} + \vec{v}_{i_3}) + \vec{v}_{i_4}$$

is equal to the square root of the sums of their lengths:

$$\sqrt{\left(\|\vec{v}_{i_2}\| + \|\vec{v}_{i_3}\|\right)^2 + \|\vec{v}_{i_4}\|^2}.$$

Since i_1 is connected to all the three other vertices, this means that there is no restriction on the relation between the vector i_1 and three other vectors. So, the maximum length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = \vec{v}_{i_1} + (\vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4}).$$

Thus, the maximum length of this sum is equal to the sum of the lengths:

$$\|\vec{v}_{i_1}\| + \sqrt{\left(\|\vec{v}_{i_2}\| + \|\vec{v}_{i_3}\|\right)^2 + \|\vec{v}_{i_4}\|^2}.$$

The use of Proposition 4.3 completes the proof.

Proof of Proposition 4.10. Vectors i_2 and i_4 are independent, so the length of the sum $\vec{v}_{i_2} + \vec{v}_{i_4}$ is equal to $\sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_4}\|^2}$. Now, there are no restrictions on the relation between $\vec{v}_{i_1}$, $\vec{v}_{i_3}$, and $\vec{v}_{i_2} + \vec{v}_{i_4}$. Thus, the maximum length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = \vec{v}_{i_1} + \vec{v}_{i_3} + (\vec{v}_{i_2} + \vec{v}_{i_4})$$

is equal to the sum of the lengths:

$$\|\vec{v}_{i_1}\| + \|\vec{v}_{i_3}\| + \sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_4}\|^2}.$$

The use of Proposition 4.3 completes the proof.

Proof of Proposition 4.11 is similar to the proofs of Propositions 5 and 6.

Proof of Proposition 4.12. The given graph means that between the four vertices, the only independent pairs are those which are not connected by an egde, i.e., pairs (i_3, i_1), (i_1, i_4), and (i_4, i_2). One can easily see that these vertices also form a linear graph. For this case, the largest value of the sum of the four vectors is computed in the proof of Proposition 4.23. The use of Proposition 4.3 completes the proof.

Proof of Proposition 4.13. Since the vertices i_1 and i_3 are not connected, this means that the measurement errors corresponding to these vertices are independent, so the length of $\vec{v}_{i_1} + \vec{v}_{i_3}$ is equal to $\sqrt{\|\vec{v}_{i_1}\|^2 + \|\vec{v}_{i_3}\|^2}$. Similarly, since the vertices i_2 and i_4 are not connected, this means that the measurement errors corresponding to these vertices are independent, so the length of $\vec{v}_{i_2} + \vec{v}_{i_4}$ is equal to $\sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_4}\|^2}$.

Each of the vertices i_1 and i_3 is connected to both i_2 and i_4, which means that we know nothing about the dependence between the corresponding measurement errors. Thus, as we have described earlier, the largest possible length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = (\vec{v}_{i_1} + \vec{v}_{i_3}) + (\vec{v}_{i_2} + \vec{v}_{i_4})$$

can be obtained by adding the lengths of $\vec{v}_{i_1} + \vec{v}_{i_3}$ and of $\vec{v}_{i_2} + \vec{v}_{i_4}$:

$$\sqrt{\|\vec{v}_{i_1}\|^2 + \|\vec{v}_{i_3}\|^2} + \sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_4}\|^2}.$$

The desired result now follows from Proposition 4.3.

Proof of Proposition 4.14. The proof of this proposition naturally follows from the geometric interpretation, in which we associate a vector to each random variable, and we

are looking for a configuration in which the sum of these vectors has the largest length. Here, the sum $\vec{v}$ of all the corresponding vectors $\vec{v} = \vec{v}_1 + \ldots + \vec{v}_n$ can be represented as

$$\sum_{i \notin V} \vec{v}_i + \vec{a},$$

where

$$\vec{a} \stackrel{\text{def}}{=} \sum_{j \in V} \vec{v}_j.$$

We do not have any restrictions on the relative orientation of the vectors $\vec{v}_i$ corresponding to "normal" errors ($i \notin V$) and of the vector $\vec{a}$. Thus, the largest possible value of the length of the sum $\vec{v}$ is equal to the sum of the lengths of the "normal" vectors $\vec{v}_i$ and of the vector $\vec{a}$:

$$\max \|\vec{v}\| = \sum_{i \notin V} \|\vec{v}_i\| + \|\vec{a}\|.$$

The values $\|\vec{v}_i\|$ are given: they are equal to $|c_i| \cdot \sigma_i$. Thus, the largest possible value of $\|\vec{v}\|$ is attained when the length $\|\vec{a}\|$ is the largest. This largest length is what in Definitions 1 and 4 we denoted by $\sigma\,(s_V, D_G, c_V)$. Thus, we get the desired formula.

The proposition is proven.

Proof of Proposition 4.16. For this graph, the value $\sigma\,(s, D_G, c)$ follows from Proposition 4.2 – it is equal to $\sqrt{c_{i_1}^2 \cdot \sigma_{i_1}^2 + c_{i_2}^2 \cdot \sigma_{i_2}^2}$. Thus, by Proposition 4.12, we get the desired result.

Proof of Proposition 4.17 is similar to the proof of Proposition 4.16.

Proof of Proposition 4.18. Since the vertices i_2 and i_3 are not connected, this means that we do not have any restrictions on the relative location of the vectors $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$, so the largest possible value of the length of the sum $\vec{v}_{i_2} + \vec{v}_{i_3}$ is equal to the sum of the lengths $\|\vec{v}_{i_2}\| + \|\vec{v}_{i_3}\|$.

The vertex i_1 is connected to both i_2 and i_3, which means that we the measurement error corresponding to i_1 is independent of the errors corresponding to i_2 and i_3. Thus, as we have described earlier, the largest possible length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} = \vec{v}_{i_1} + (\vec{v}_{i_2} + \vec{v}_{i_3})$$

is equal to

$$\sqrt{\|\vec{v}_{i_1}\|^2 + (\|\vec{v}_{i_2}\| + \|\vec{v}_{i_3}\|)^2}.$$

The desired result now follows from Proposition 4.14.

Proof of Proposition 4.19 is similar to the proof of Proposition 4.18.

Proof of Proposition 4.20. The vectors $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$ are independent, so the length of their sum $\vec{v}_{i_2} + \vec{v}_{i_3}$ is equal to $\sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2}$. There is no edge between i_4 and the group of vertices (i_2, i_3), this means there is no restriction on the relation between $\vec{v}_{i_4}$ and $\vec{v}_{i_2} + \vec{v}_{i_3}$. Thus, the largest possible length of the sum

$$\vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = (\vec{v}_{i_2} + \vec{v}_{i_3}) + \vec{v}_{i_4}$$

is equal to the sum of their lengths:

$$\sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2} + \|\vec{v}_{i_4}\|.$$

Since i_1 is connected to all the three other vertices, this means that this vector is orthogonal to three other vectors – and thus, to their sum. So, the maximum length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = \vec{v}_{i_1} + (\vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4})\,.$$

is equal to

$$\sqrt{\|\vec{v}_{i_1}\|^2 + \left(\sqrt{\|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2} + \|\vec{v}_{i_4}\|\right)^2}.$$

The use of Proposition 4.11 completes the proof.

Proof of Proposition 4.21. There is no constraint on the vectors i_2 and i_4, so the maximum length of the sum $\vec{v}_{i_2} + \vec{v}_{i_4}$ is equal to the sum of their length: $\|\vec{v}_{i_2}\| + \|\vec{v}_{i_4}\|$. Now, the three vectors $\vec{v}_{i_1}$, $\vec{v}_{i_3}$, and $\vec{v}_{i_2} + \vec{v}_{i_4}$ are independent. Thus, the maximum length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = \vec{v}_{i_1} + \vec{v}_{i_3} + (\vec{v}_{i_2} + \vec{v}_{i_4})$$

is equal to:

$$\sqrt{\|\vec{v}_{i_1}\|^2 + \|\vec{v}_{i_3}\|^2 + (\|\vec{v}_{i_2}\| + \|\vec{v}_{i_4}\|)^2}.$$

The use of Proposition 4.11 completes the proof.

Proof of Proposition 4.22 is similar to the proof of Propositions 16 and 17.

Proof of Proposition 4.23. In accordance with Proposition 4.14, we need to compute $\sigma \stackrel{\text{def}}{=} \sigma(s_V, D_G, c_V)$, the largest possible length of the vector $\vec{v} = \vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4}$. In general, the square $\|\vec{v}\|^2$ of the length $\|\vec{v}\|$ of the sum $\vec{v} = \vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4}$ of the four vectors is equal to

$$\|\vec{v}\|^2 = \|\vec{v}_{i_1}\|^2 + \|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2 + \|\vec{v}_{i_4}\|^2 +$$

$$2\vec{v}_{i_1} \cdot \vec{v}_{i_2} + 2\vec{v}_{i_1} \cdot \vec{v}_{i_3} + 2\vec{v}_{i_1} \cdot \vec{v}_{i_4} + 2\vec{v}_{i_2} \cdot \vec{v}_{i_3} + 2\vec{v}_{i_2} \cdot \vec{v}_{i_4} + 2\vec{v}_{i_3} \cdot \vec{v}_{i_4}.$$

For the situation described by the given graph, vector $\vec{v}_{i_1}$ is orthogonal to $\vec{v}_{i_2}$, the vector $\vec{v}_{i_2}$ is orthogonal to $\vec{v}_{i_3}$, and the vector $\vec{v}_{i_3}$ is orthogonal to $\vec{v}_{i_4}$. Thus, we have

$$\|\vec{v}\|^2 = \|\vec{v}_{i_1}\|^2 + \|\vec{v}_{i_2}\|^2 + \|\vec{v}_{i_3}\|^2 + \|\vec{v}_{i_4}\|^2 + 2\vec{v}_{i_1} \cdot \vec{v}_{i_3} + 2\vec{v}_{i_1} \cdot \vec{v}_{i_4} + 2\vec{v}_{i_2} \cdot \vec{v}_{i_4}.$$

The length of each vector $\|\vec{v}_{i_j}\|$ is fixed $\|\vec{v}_{i_j}\|^2 = v_{i_j}^2$, so to maximize the length of the sum, we need to maximize the sum of the remaining terms:

$$2\vec{v}_{i_1} \cdot \vec{v}_{i_3} + 2\vec{v}_{i_1} \cdot \vec{v}_{i_4} + 2\vec{v}_{i_2} \cdot \vec{v}_{i_4}.$$

Let us denote the half of this sum by J, then the sum itself becomes equal to $2J$.

We need to maximize the sum $2J$ under the constraints

$$\|\vec{v}_{i_j}\|^2 = v_{i_j}^2 \text{ for all } j, \vec{v}_{i_1} \cdot \vec{v}_{i_2} = 0, \vec{v}_{i_2} \cdot \vec{v}_{i_3} = 0, \text{ and } \vec{v}_{i_3} \cdot \vec{v}_{i_4} = 0.$$

By using the Lagrange multiplier method, we can reduce the above-described conditional optimization problem to the following unconstrained optimization problem:

$$2\vec{v}_{i_1} \cdot \vec{v}_{i_3} + 2\vec{v}_{i_1} \cdot \vec{v}_{i_4} + 2\vec{v}_{i_2} \cdot \vec{v}_{i_4} + \sum_{j=1}^{4} \lambda_j \cdot \|\vec{v}_{i_j}\|^2 + \sum_{j=1}^{3} \mu_j \cdot \vec{v}_{i_j} \cdot \vec{v}_{i_{j+1}},$$

where λ_j and μ_j are Lagrange multipliers.

Differentiating this expression with respect to $\vec{v}_{i_2}$ and equating the derivative to 0, we conclude that

$$2\vec{v}_{i_4} + 2\lambda_2 \cdot \vec{v}_{i_2} + \mu_1 \cdot \vec{v}_{i_1} + \mu_2 \cdot \vec{v}_{i_3} = 0,$$

hence

$$\vec{v}_{i_4} = -\lambda_2 \cdot \vec{v}_{i_2} - \frac{1}{2} \cdot \mu_1 \cdot \vec{v}_{i_1} - \frac{1}{2} \cdot \mu_2 \cdot \vec{v}_{i_3}.$$

So, the vector $\vec{v}_{i_4}$ belongs to the linear space generated by vectors $\vec{v}_{i_2}$, $\vec{v}_{i_3}$, and $\vec{v}_{i_1}$. Let us denote the unit vectors in the directions of $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$ by, correspondingly,

$$\vec{e}_2 = \frac{\vec{v}_{i_2}}{v_{i_2}} \text{ and } \vec{e}_3 = \frac{\vec{v}_{i_3}}{v_{i_3}}.$$

Since the vectors $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$ are orthogonal, the unit vectors $\vec{e}_2$ and $\vec{e}_3$ are orthogonal too, so they can be viewed as two vectors from the orthonormal basis in the linear space generated by the vectors $\vec{v}_{i_2}$, $\vec{v}_{i_3}$, and $\vec{v}_{i_1}$.

- If this linear space is 3-dimensional, in this 3-D spoace we can select the third unit vector $\vec{e}$ which is orthogonal to both $\vec{e}_2$ and $\vec{e}_3$.
- If the above linear space is 2-dimensional – i.e., if $\vec{v}_{i_1}$ lies in the 2-D space generated by $\vec{v}_{i_2}$ and $\vec{v}_{i_3}$ – then let us take, as $\vec{e}$, any unit vector which is orthogonal to both $\vec{e}_2$ and $\vec{e}_3$.

In both cases, vectors $\vec{v}_{i_1}$ and $\vec{v}_{i_4}$ belong to the linear space generated by the vectors $\vec{e}_2$, $\vec{e}_3$, and $\vec{e}$. In particular, this means that $\vec{v}_{i_1} = c_{12} \cdot \vec{e}_2 + c_{13} \cdot \vec{e}_3 + c_1 \cdot \vec{e}$ for some numbers c_{12}, c_{13}, and c_1. Since $\vec{v}_{i_1} \perp \vec{v}_{i_2}$, we have $c_{12} = 0$, so $\vec{v}_{i_1} = c_{13} \cdot \vec{e}_3 + c_1 \cdot \vec{e}$. From this formula, we conclude that $v_{i_1}^2 = \|\vec{v}_{i_1}\|^2 = c_{13}^2 + c_1^2$, so $c_1^2 \le v_{i_1}^2$. Let us denote the ratio c_1/v_{i_1} by β_1, then $c_1 = v_{i_1} \cdot \beta_1$ and, correspondingly, $c_{13} = v_{i_1} \cdot \sqrt{1 - \beta_1^2}$. So, the expression for $\vec{v}_{i_1}$ takes the form

$$\vec{v}_{i_1} = v_{i_1} \cdot \sqrt{1 - \beta_1^2} \cdot \vec{e}_3 + v_{i_1} \cdot \beta_1 \cdot \vec{e}.$$

Similarly, we can conclude that

$$\vec{v}_{i_4} = v_{i_4} \cdot \sqrt{1 - \beta_4^2} \cdot \vec{e}_2 + v_{i_4} \cdot \beta_4 \cdot \vec{e},$$

for some value β_4 for which $|\beta_4| \le 1$. For each pair of orthogonal vectors $\vec{e}_2$ and $\vec{v}_{i_3}$ of lengths v_{i_2} and v_{i_3}, the above-defined vectors satisfy all the constraints. So, what remains

is to find the values β_1 and β_4 for which the expression

$$2\vec{v}_{i_1} \cdot \vec{v}_{i_3} + 2\vec{v}_{i_1} \cdot \vec{v}_{i_4} + 2\vec{v}_{i_2} \cdot \vec{v}_{i_4}$$

attaints its largest value – i.e., equivalently, for which the above-defined half-of-the-maximized-expression

$$J = \vec{v}_{i_1} \cdot \vec{v}_{i_3} + \vec{v}_{i_1} \cdot \vec{v}_{i_4} + \vec{v}_{i_2} \cdot \vec{v}_{i_4}$$

attains its largest value. Substituting the above expressions for $\vec{v}_{i_1}$ and $\vec{v}_{i_4}$ into this formula, and taking into account that, by our choice of $\vec{e}_2$ and $\vec{e}_3$, we have $\vec{v}_{i_2} = v_{i_2} \cdot \vec{e}_2$ and $\vec{v}_{i_3} = v_{i_3} \cdot \vec{e}_3$, we conclude that

$$J = v_{i_1} \cdot v_{i_3} \cdot \sqrt{1 - \beta_1^2} + v_{i_2} \cdot v_{i_4} \cdot \sqrt{1 - \beta_4^2} + \beta_1 \cdot \beta_4 \cdot v_{i_1} \cdot v_{i_4}.$$

Each of the unknown β_1 and β_4 has values from the interval $[-1, 1]$. Thus, for each of the variables β_1 and β_4, the maximum of this expression is attained:

- either at one of the endpoints -1 and 1 of this interval,
- or at the point inside this interval, in which case the derivative with respect to this variable should be equal to 0.

We have two cases for each of the two variables β_1 and β_4, so overall, we need to consider all $2 \cdot 2 = 4$ cases. To find the largest possible value of the expression J, we need to consider all four possible cases, and find the largest of the corresponding values. Let us consider these cases one by one.

Case 1. If both values β_1 and β_4 are equal to ± 1, then we get $J = \pm v_{i_1} \cdot v_{i_4}$. The largest of these values is when the sign is positive, then the value of the quantity J is equal to $J_1 = v_{i_1} \cdot v_{i_4}$.

Case 2. Let us now consider the case when $\beta_1 = \pm 1$ and $\beta_4 \in (-1, 1)$. In this case, the expression J takes the form $J = v_{i_2} \cdot v_{i_4} \cdot \sqrt{1 - \beta_4^2} \pm v_{i_1} \cdot v_{i_4} \cdot \beta_4$. Differentiating this expression with respect to β_4 and equating the derivative to 0, we get

$$-\frac{2\beta_4 \cdot v_{i_2} \cdot v_{i_4}}{2\sqrt{1 - \beta_4^2}} \pm v_{i_1} \cdot v_{i_4} \cdot \beta_1 = 0.$$

If we divide both sides by v_{i_4}, divide both the numerator and the denominator of the fraction by a common factor 2, and multiply both sides by the denominator, we get

$$\beta_4 \cdot v_{i_2} = \pm\sqrt{1-\beta_4^2} \cdot v_{i_1}.$$

If we square both sides, we get

$$\beta_4^2 \cdot v_{i_2}^2 = \left(1-\beta_4^2\right) \cdot v_{i_1}^2 = v_{i_1}^2 - \beta_4^2 \cdot v_{i_1}^2.$$

So

$$\beta_4^2 \cdot \left(v_{i_1}^2 + v_{i_2}^2\right) = v_{i_2}^2$$

and

$$\beta_4^2 = \frac{v_{i_2}^2}{v_{i_1}^2 + v_{i_2}^2}.$$

Therefore,

$$1-\beta_4^2 = \frac{v_{i_1}^2}{v_{i_1}^2 + v_{i_2}^2},$$

so

$$\beta_4 = \pm\frac{v_{i_2}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}}$$

and

$$\sqrt{1-\beta_4^2} = \pm\frac{v_{i_1}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}}.$$

Substituting these expressions into the formula for J, we conclude that

$$J = \pm\frac{v_{i_2}^2 \cdot v_{i_4}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}} \pm \frac{v_{i_1}^2 \cdot v_{i_4}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}}.$$

The largest value of this expression is attained when both signs are positive, so we get

$$J = \frac{v_{i_2}^2 \cdot v_{i_4}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}} \pm \frac{v_{i_1}^2 \cdot v_{i_4}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}} = \frac{v_{i_4} \cdot \left(v_{i_1}^2 + v_{i_2}^2\right)}{\sqrt{v_{i_1}^2 + v_{i_2}^2}}$$

and thus, the value J is equal to

$$J_2 = v_{i_4} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}.$$

In this case, the largest value of $\|\vec{v}\|^2$ is equal to:

$$\sigma_2^2 = v_{i_1}^2 + v_{i_2}^2 + v_{i_3}^2 + v_{i_4}^2 + 2v_{i_4} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2} =$$

$$v_{i_3}^2 + \left(\left(v_{i_1}^2 + v_{i_2}^2 \right) + v_{i_4}^2 + 2v_{i_4} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2} \right) =$$

$$v_{i_3}^2 + \left(\sqrt{v_{i_1}^2 + v_{i_2}^2} + v_{i_4} \right)^2.$$

Comparing Case 1 and Case 2. Since $v_{i_1}^2 + v_{i_2}^2 > v_{i_1}^2$, we have $\sqrt{v_{i_1}^2 + v_{i_2}^2} > v_{i_1}$, hence $J_2 = v_{i_4} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2} > v_{i_1} \cdot v_{i_4} = J_1$. Thus, when we are looking for the largest value of the expression J, we can safely ignore Case 1, since the values obtained in Case 2 can be larger than anything we get in Case 1.

Case 3. Similarly, we can consider the case when $\beta_4 = \pm 1$ and $\beta_1 \in (-1, 1)$. In this case, we get the largest possible value of J equal to $J_3 = v_{i_1} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2}$, so the largest possible value of σ^2 is equal to:

$$\sigma_3^2 = v_{i_2}^2 + \left(v_{i_1} + \sqrt{v_{i_3}^2 + v_{i_4}^2} \right).$$

Case 4. Finally, let us consider the case when for the pair (β_1, β_4) at which the expression J attains its largest value, both values β_1 and β_4 are located inside the interval $(-1, 1)$. In this case, to find the maximum of the expression J, we need to differentiate it with respect to the unknowns β_1 and β_4 and equate the resulting derivatives to 0. If we differentiate by β_1, we get

$$-\frac{2\beta_1 \cdot v_{i_1} \cdot v_{i_3}}{2\sqrt{1 - \beta_1^2}} + v_{i_1} \cdot v_{i_4} \cdot \beta_4 = 0.$$

Thus,

$$\beta_4 = \frac{\beta_1 \cdot v_{i_3}}{\sqrt{1 - \beta_1^2} \cdot v_{i_4}},$$

and

$$\beta_4^2 = \frac{\beta_1^2 \cdot v_{i_3}^2}{(1 - \beta_1^2) \cdot v_{i_4}^2}.$$

Differentiating the above expression for J with respect to β_4 and equating the derivative to 0, we conclude that

$$-\frac{2\beta_4 \cdot v_{i_2} \cdot v_{i_4}}{2\sqrt{1-\beta_4^2}} + v_{i_1} \cdot v_{i_4} \cdot \beta_1 = 0.$$

If we divide both sides by v_{i_4}, divide both the numerator and the denominator of the fraction by a common factor 2, and multiply both sides by the denominator, we get

$$\beta_4 \cdot v_{i_2} = \sqrt{1-\beta_4^2} \cdot v_{i_1} \cdot \beta_1.$$

If we square both sides, we get

$$\beta_4^2 \cdot v_{i_2}^2 = \left(1-\beta_4^2\right) \cdot v_{i_1}^2 \cdot \beta_1^2 = v_{i_1}^2 \cdot \beta_1^2 - \beta_4^2 \cdot v_{i_1}^2 \cdot \beta_1^2.$$

Substituting the above expression for β_4^2 into this formula, we get

$$\frac{\beta_1^2 \cdot v_{i_2}^2 \cdot v_2^3}{\sqrt{1-\beta_1^2 \cdot v_{i_4}^2}} = \beta_1^2 \cdot v_{i_1}^2 - \frac{\beta_1^4 \cdot v_{i_1}^2 \cdot v_{i_3}^2}{\sqrt{1-\beta_1^2 \cdot v_{i_4}^2}}.$$

Case 4, subcase when $\beta_1 = 0$. Both sides of this equality contain the common factor β_1. So, it is possible that $\beta_1 = 0$, in which case $\beta_4 = 0$, and the expression J attains the value

$$J_0 = v_{i_1} \cdot v_{i_3} + v_{i_2} \cdot v_{i_4}.$$

In this case, the value of σ^2 is equal to:

$$\sigma_0^2 = v_{i_1}^2 + v_{i_2}^2 + v_{i_3}^2 + v_{i_4}^2 + 2v_{i_1} \cdot v_{i_3} + 2v_{i_2} \cdot v_{i_4} =$$

$$\left(v_{i_1}^2 + v_{i_3}^2 + 2v_{i_1} \cdot v_{i_3}\right) + \left(v_{i_2}^2 + v_{i_4}^2 + 2v_{i_2} \cdot v_{i_4}\right) =$$

$$\left(v_{i_1} + v_{i_3}\right)^2 + \left(v_{i_2} + v_{i_4}\right)^2.$$

Case 4, subcase when $\beta_1 \neq 0$. If $\beta_1 \neq 0$, then we can divide both sides of the above equality by β_1^2. Multiplying both sides by the denominator, we get

$$v_{i_2}^2 \cdot v_{i_3}^2 = v_{i_1}^2 \cdot v_{i_4}^2 \cdot \left(1-\beta_1^2\right) - \beta_1^2 \cdot v_{i_1}^2 \cdot v_{i_3}^2,$$

so

$$v_{i_2}^2 \cdot v_{i_3}^2 = v_{i_1}^2 \cdot v_{i_4}^2 - \beta_1^2 \cdot v_{i_1}^2 \cdot v_{i_4}^2 - \beta_1^2 \cdot v_{i_1}^2 \cdot v_{i_3}^2.$$

If we move all the terms containing β_1^2 to the left-hand side and all the other terms to the right-hand side, we get:

$$\beta_1 v_{i_1}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right) = v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_1}^2 \cdot v_{i_3}^2,$$

thus

$$\beta_1^2 = \frac{v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2}{v_{i_1}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right)}.$$

Here:

- when $v_{i_1} \cdot v_{i_4} < v_{i_2} \cdot v_{i_3}$, the right-hand side is negative, so we cannot have such a case;
- when $v_{i_1} \cdot v_{i_4} = v_{i_2} \cdot v_{i_3}$, then $\beta_1 = 0$, and we have already analyzed this case.

So, the only possibility to have $\beta_1 \neq 0$ is when $v_{i_1} \cdot v_{i_4} > v_{i_2} \cdot v_{i_3}$.

In general, the situation does not change if we swap 1 and 4 and swap 2 and 3. Thus, for β_4^2, we get a similar expression

$$\beta_4^2 = \frac{v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2}{v_{i_4}^2 \cdot \left(v_{i_1}^2 + v_{i_2}^2\right)}.$$

From the expressions for β_1^2 and β_4^2, we conclude that

$$1 - \beta_1^2 = 1 - \frac{v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2}{v_{i_1}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right)} =$$

$$\frac{v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_1}^2 \cdot v_{i_4}^2 + v_{i_2}^2 \cdot v_{i_3}^2}{v_{i_1}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right)} =$$

$$\frac{v_{i_3}^2 \cdot \left(v_{i_1}^2 + v_{i_2}^2\right)}{v_{i_1}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right)}.$$

Similarly, we have

$$1 - \beta_4^2 = \frac{v_{i_2}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right)}{v_{i_4}^2 \cdot \left(v_{i_1}^2 + v_{i_2}^2\right)}.$$

Thus, for the expression J, we get the value

$$J_4 = v_{i_1} \cdot v_{i_3} \cdot \frac{v_{i_3} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}}{v_{i_1} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2}} + v_{i_2} \cdot v_{i_4} \cdot \frac{v_{i_2} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2}}{v_{i_4} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}} +$$

$$v_{i_1} \cdot v_{i_4} \cdot \frac{v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2}{v_{i_1} \cdot v_{i_4} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}}.$$

We can somewhat simplify this expression if:

- in the first term, we delete v_{i_1} in the numerator and in the denominator,
- in the second term, we delete v_{i_4} from the numerator and from the denominator, and
- in the third term, we delete both v_{i_1} and v_{i_4} from the numerator and from the denominator.

Then, we get:

$$J_4 = v_{i_3} \cdot \frac{v_{i_3} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}}{\sqrt{v_{i_3}^2 + v_{i_4}^2}} + v_{i_2} \cdot \frac{v_{i_2} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2}}{\sqrt{v_{i_1}^2 + v_{i_2}^2}} + \frac{v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2}{\sqrt{v_{i_3}^2 + v_{i_4}^2} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}}.$$

If we bring all the terms to the common denominator $\sqrt{v_{i_3}^2 + v_{i_4}^2} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}$, then we get

$$J_4 = \frac{v_{i_3}^2 \cdot \left(v_{i_1}^2 + v_{i_2}^2\right) + v_{i_2}^2 \cdot \left(v_{i_3}^2 + v_{i_4}^2\right) + v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2}{\sqrt{v_{i_3}^2 + v_{i_4}^2} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}}.$$

The numerator of this expression has the form

$$v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_4}^2 + v_{i_1}^2 \cdot v_{i_4}^2 - v_{i_2}^2 \cdot v_{i_3}^2 =$$

$$v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_4}^2 + v_{i_1}^2 \cdot v_{i_4}^2 =$$

$$\left(v_{i_1}^2 + v_{i_2}^2\right) \cdot \left(v_{i_3}^2 + v_{i_4}^2\right).$$

Thus, we get

$$J_4 = \frac{\left(v_{i_1}^2 + v_{i_2}^2\right) \cdot \left(v_{i_3}^2 + v_{i_4}^2\right)}{\sqrt{v_{i_3}^2 + v_{i_4}^2} \cdot \sqrt{v_{i_1}^2 + v_{i_2}^2}},$$

i.e.,

$$J_4 = \sqrt{v_{i_1}^2 + v_{i_2}^2} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2}.$$

Comparing J_4 with J_2 and J_3. One can easily see that we always have $J_2^2 \le J_4^2$ and $J_3^2 \le J_4^2$, thus $J_2 \le J_4$ and $J_3 \le J_4$. Thus, if the estimate J_4 is possible, there is no need to consider J_2 and J_3, we only need to cosider J_4 and J_0.

Comparing J_4 and J_0. Let us show that we always have $J_0 \le J_4$, i.e.,

$$v_{i_1} \cdot v_{i_3} + v_{i_2} \cdot v_{i_4} \le \sqrt{v_{i_1}^2 + v_{i_2}^2} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2}.$$

Indeed, this inequality between positive numbers is equivalent to a similar inequality between their squares:

$$v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_4}^2 + 2v_{i_1} \cdot v_{i_2} \cdot v_{i_3} \cdot v_{i_4} \le \left(v_{i_1}^2 + v_{i_2}^2\right) \cdot \left(v_{i_3}^2 + v_{i_4}^2\right),$$

i.e.,

$$v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_4}^2 + 2v_{i_1} \cdot v_{i_2} \cdot v_{i_3} \cdot v_{i_4} \le v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_1}^2 \cdot v_{i_4}^2 + v_{i_2}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_4}^2.$$

Subtracting $v_{i_1}^2 \cdot v_{i_3}^2 + v_{i_2}^2 \cdot v_{i_4}^2$ from both sides of this inequality, we get an equivalent inequality

$$2v_{i_1} \cdot v_{i_2} \cdot v_{i_3} \cdot v_{i_4} \le v_{i_1}^2 \cdot v_{i_4}^2 + v_{i_2}^2 \cdot v_{i_3}^2,$$

i.e., equivalently,

$$0 \le v_{i_1}^2 \cdot v_{i_4}^2 + v_{i_2}^2 \cdot v_{i_3}^2 - 2v_{i_1} \cdot v_{i_2} \cdot v_{i_3} \cdot v_{i_4} = \left(v_{i_1} \cdot v_{i_4} - v_{i_2} \cdot v_{i_3}\right)^2,$$

which is, of course, always true. Thus, when the estimate J_4 is possible, we do not need to consider the value J_0 either: it is sufficient to take $J = J_4$.

Value of σ in case J_4 is possible: conclusion. So, if the value J_4 is possible, we get

$$\sigma^2 = v_{i_1}^2 + v_{i_2}^2 + v_{i_3}^2 + v_{i_4}^2 + 2\sqrt{v_{i_1}^2 + v_{i_2}^2} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2} =$$

$$\left(v_{i_1}^2 + v_{i_2}^2\right) + \left(v_{i_3}^2 + v_{i_4}^2\right) + 2\sqrt{v_{i_1}^2 + v_{i_2}^2} \cdot \sqrt{v_{i_3}^2 + v_{i_4}^2} = \left(\sqrt{v_{i_1}^2 + v_{i_2}^2} + \sqrt{v_{i_3}^2 + v_{i_4}^2}\right)^2,$$

so $\sigma = \sqrt{v_{i_1}^2 + v_{i_2}^2} + \sqrt{v_{i_3}^2 + v_{i_4}^2}$.

General comment. The desired result for this case now follows from Proposition 4.14.

Proof of Proposition 4.24. Since the vertices i_2 and i_4 are not connected, this means that we do not have any restrictions on the relative location of the vectors $\vec{v}_{i_2}$ and $\vec{v}_{i_4}$, so the largest possible value of the length of the sum $\vec{v}_{i_2} + \vec{v}_{i_4}$ is equal to the sum of the lengths $\|\vec{v}_{i_2}\| + \|\vec{v}_{i_4}\|$. Similarly, since the vertices i_1 and i_3 are not connected, this means that we do not have any restrictions on the relative location of the vectors $\vec{v}_{i_1}$ and $\vec{v}_{i_3}$, so the largest possible value of the length of the sum $\vec{v}_{i_1} + \vec{v}_{i_3}$ is equal to the sum of the lengths $\|\vec{v}_{i_1}\| + \|\vec{v}_{i_3}\|$.

Each of the vertices i_1 and i_3 is connected to both i_2 and i_4, which means that the measurement errors corresponding to i_1 and i_3 are independent of the errors corresponding to i_2 and i_4. Thus, as we have described earlier, the largest possible length of the sum

$$\vec{v}_{i_1} + \vec{v}_{i_2} + \vec{v}_{i_3} + \vec{v}_{i_4} = (\vec{v}_{i_1} + \vec{v}_{i_3}) + (\vec{v}_{i_2} + \vec{v}_{i_4})$$

is equal to

$$\sqrt{\left(\|\vec{v}_{i_1}\| + \|\vec{v}_{i_3}\|\right)^2 + \left(\|\vec{v}_{i_2}\| + \|\vec{v}_{i_3}\|\right)^2}.$$

The desired result now follows from Proposition 4.14.

4.6 Conclusions and Recommendations for Future Work

Conclusions. In this chapter, we deal with the third of the four challenges of practical computer-enhanced measurements. This challenge is related to the fact that in the past, when we could only afford a few measurements, these measurements were usually performed

by similar measuring instruments, instruments for which we had a good understanding of what causes their measurement errors.

- In some situations, most measurement errors were causes by internal features of the instruments. In this case, the corresponding measurement errors were independent.
- In other situations, mostly external features were dominant, in which case we do have any information about the relation between different measurement errors.

In both types of situations, formulas were developed for processing the resulting uncertainty. With the possibility to perform numerous measurements and process their results, we often encounter situations when some pairs of measurement errors are independent but for other pairs of measurement errors, we do not have any information about their relation.

In this chapter, our objective was to come up with techniques for processing measurement results in situations which are slightly different from the above-described well-studied ones; namely:

- for the situations when for most pairs of measuring instruments, we know that the corresponding measuring errors are independent, but for a few pairs, we do not have any information about their dependence, and
- for the situations in which for most pairs of measuring instruments, we have no information about the dependence between the corresponding measurement errors, but for some pairs, we know that the corresponding measurement errors are independent.

As a result of our analysis, we provide new explicit easy-to-implement formulas describing the uncertainty of the result of data processing in above-described situations.

Recommendations for future work. In the current chapter, we only deal with the cases when for the most pairs, we have information of the same type, and only for a small number of pairs, we have different information. It is necessary to extend our analysis to situations when we have a larger number of pairs with different information.

Chapter 5

Fault Detection in a Smart Electric Grid: Geometric Analysis

In this chapter, we deal with the last of the four challenges of practical computer-enhanced measurements – the challenge related to the need to extract information from the measurement results. Specifically, we deal with the simplest case when we only know the ordering of the measurement results, but not the actual numerical values.

An important example of such a situation is locating fault in a smart electric grid. The main idea behind a smart grid is to equip the grid with a dense lattice of sensors monitoring the state of the grid. If there is a fault, the sensors closer to the fault will detect larger deviations from the normal readings than sensors that are farther away. In this chapter, we show that this fact can be used to locate the fault with high accuracy.

5.1 What Is a Smart Electric Grid

The main idea is to set up a lattice of sensors that would monitor the electric grid; see, e.g., [17]. Based on the measurement results provided by the sensors:

- we would get a good picture of the current state of the grid, and
- we would be able to effectively control it.

5.2 How the Grid of Sensors Can Detect Faults

Each sensor measures characteristics of the electric current at its location. Each fault affects all the sensors, some more, some less.

By observing the changes in the sensor signals, we can detect the existence of the fault. We can also get some information about the fault's location.

Sensors closer to the fault's location will detect a stronger change in their measurement results than sensors which are further away. Thus, by comparing the measurement results of the two sensors, we can decide whether the fault is:

- closer to the first sensor or
- closer to the second sensor.

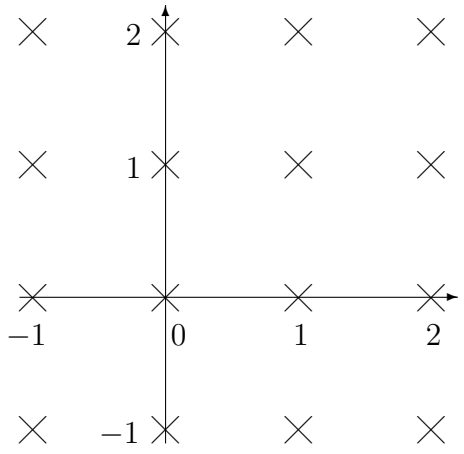
5.3 Let Us Describe This Situation in Precise Terms

Let us consider the case when the sensors form a (potentially infinite) rectangular lattice. For simplicity of analysis, let us select a coordinate system in which:

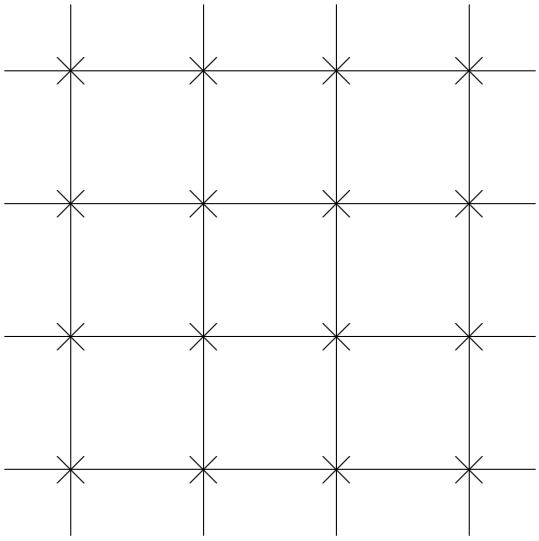
- the location of one the sensors is the starting point $(0, 0)$, and

- the distance between the closest sensors is used as a measuring unit.

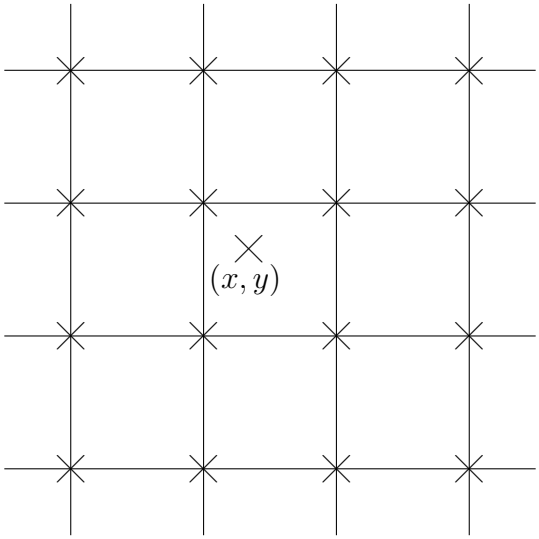
In this coordinate system, sensors are located at all the points (a, b) with integer coordinates.



These sensors divide the plane into squares $[a, a + 1] \times [b, b + 1]$.



Each spatial location (x, y) is in one of these squares:

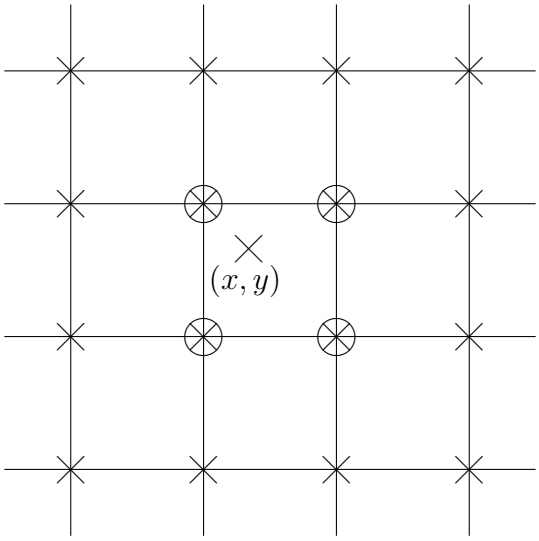


One can easily check that:

- for each spatial location within a square,
- the vertices (a, b), $(a, b+1)$, $(a+1, b)$, and $(a+1, b+1)$ of this square are the closest grid points.

Thus:

- by finding the 4 sensors at which the disturbance signal is the strongest,
- we can find the square that contains the location of the fault.



5.4 Research Question

Can we determine the location of the fault more accurately than "somewhere in the square"?

5.5 Our Answer

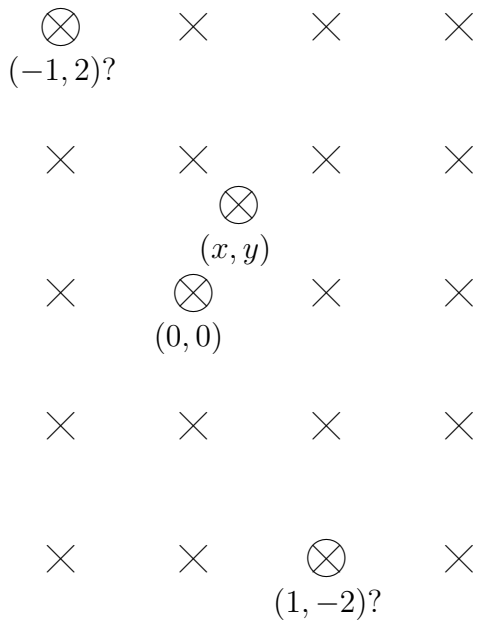
We show that, in principle:

- by using the lattice of sensors,
- we can locate the fault with any desired accuracy.

Indeed, without losing generality, let us assume that the square containing the fault is the square $[0,1] \times [0,1]$. In other words, we know that the coordinates (x, y) of the fault satisfy the inequalities $0 \le x \le 1$ and $0 \le y \le 1$.

For each pair of positive integers (p, q), we can check whether

- the sensor at $(p, -q)$ gets a stronger signal than
- the sensor at $(-p, q)$.



The first sensor's signal is stronger if and only if:

- the squared distance $d^2(f, s_1) = (x - p)^2 + (y - (-q))^2$ between the fault f and the first sensor s_1 is smaller than
- the squared distance $d^2(f, s_2) = (x - (-p))^2 + (y - q)^2$ to the second sensor.

One can check that $d^2(f, s_1) < d^2(f, s_2)$ if and only if $q \cdot y < p \cdot x$, i.e., if and only if

$$\frac{y}{x} < \frac{p}{q}.$$

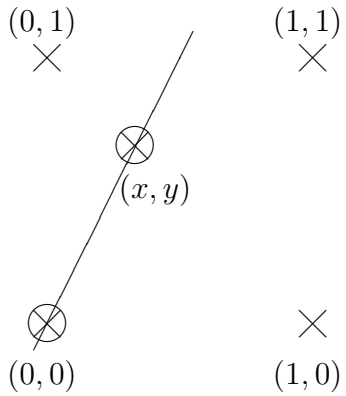
A real number can be uniquely determined if we know:

- which rational numbers p/q are smaller than this number and
- which are larger.

Thus:

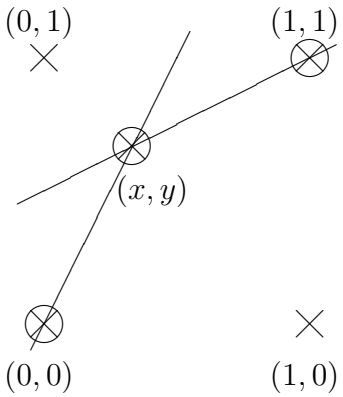
- by comparing signals from different sensors,
- we can determine the ratio $r \stackrel{\text{def}}{=} y/x$ with any given accuracy.

Hence, we can determine the line $y = r \cdot x$ going through $(0, 0)$ that contains the fault:



Similarly, we can find a straight line going through the point $(1, 1)$ that contains the fault. Thus:

- the fault's location can be uniquely determined
- as the intersection of these two straight lines.



5.6 Conclusions and Recommendations for Future Work

Conclusions. In this chapter, we deal with the last of the four challenges of practical computer-enhanced measurements. This challenge is related to the need to extract useful information from the measurement results. In the current chapter, we analyze the simplest case of this challenge, when we only know the ordering of the measurement results, but not the actual numerical values.

As a result of our analysis, we come up with a theoretical result explaining – on the example of fault location in an electric grid – that information about the ordering of measurement results can be sufficient to accurately locate the fault.

Recommendations for future work. In the current chapter, we only deal with the case when we know the ordering of the measurement results, but not the numerical values themselves. It is necessary to extend our analysis to situations when we have (and can use) numerical values as well.

Chapter 6

Conclusions and Recommendations for Future Work

6.1 Conclusions

Challenges. In this thesis, we dealt with the main challenges related to practical computer-enhanced measurements:

- The first challenge is related to the fact that the existing metrological recommendations are mostly based on the previous practice, when we could only afford to have a small number of measurements. As a result, the same system that in the past (when fewer measurements were possible) would have successfully passed the metrological analysis is no longer certified when more measurement results are available. This is a serious problem that, e.g., halted the design of the International Thermonuclear Experimental Reactor ITER.

- The second challenge is related to the fact that in the past, when there were few affordable measuring instruments and we could only afford a few measurements, there were not that many options. In such cases, planning measurements simply meant selecting one of these options. So, we could plan the measurements "by hand". Nowadays, with a potential to perform a large number of measurements and the availability of many different measuring instruments, the number of possible measurement options becomes so large that we need to develop methods for optimal planning. There exist techniques for such planning, but these techniques are mostly based on limited num-

ber of measurements. For situations when we have a large number of measurements, to the best of our knowledge, no practical general methods are known – even for the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results.

- The third challenge is related to the fact that in the past, when we could only afford a few measurements, these measurements were usually performed by similar measuring instruments, instruments for which we had a good understanding of what causes their measurement errors. In some situations, most measurement errors were causes by internal features of the instruments. In this case, the corresponding measurement errors were independent. In other situations, mostly external features were dominant, in which case we do have any information about the relation between different measurement errors. In both types of situations, formulas were developed for processing the resulting uncertainty. With the possibility to perform numerous measurements and process their results, we often encounter situations when some pairs of measurement errors are independent but for other pairs of measurement errors, we do not have any information about their relation.

- The final – fourth – challenge is how to extract useful information from all these measurement results.

Objectives. In this thesis, our main objective was to deal with the simplest possible cases of these challenges:

- In relation to the first challenge, we analyzed how to make sure that the measurement standards do not lead to the current counterintuitive practice of reducing the number of measurements.

- In relation to the second challenge, we analyzed how to come up with optimal experiment design for the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results.

- In relation to the third challenge, we analyzed how to come up with techniques for processing measurement results in situations which are slightly different from the above-described well-studied ones; namely:
  - for the situations when for most pairs of measuring instruments, we know that the corresponding measuring errors are independent, but for a few pairs, we do not have any information about their dependence, and
  - for the situations in which for most pairs of measuring instruments, we have no information about the dependence between the corresponding measurement errors, but for some pairs, we know that the corresponding measurement errors are independent.
- In relation to the fourth challenge, we analyzed how to extract information from the measurements, in the simplest case when we only know the ordering of the measurement results, but not the actual numerical values.

Contributions. As a result of our analysis, we came up with the following contributions:

- For the first objective, we propose the idea of how to change the standards, so as to avoid the above-mentioned unfortunate situations, when additional measurements can (and do) put the system at risk of not being approved.
- For the second objective, we provide a theoretical analysis of the problem and find a new explicit formulas for the optimal measurement design. As an interesting side effect of this theoretical analysis, we come up with an explanation of why measurement accuracy is usually described by listing absolute and relative error components. To the best of our knowledge, ours is the first theoretical explanation for this widely used practice.
- For the third challenge, we provide new explicit easy-to-implement formulas describing the uncertainty of the result of data processing in above-described situations.

- Finally, for the fourth challenge, we provide a theoretical result explaining – on the example of fault location in an electric grid – that information about the ordering of measurement results can be sufficient to accurately locate the fault.

6.2 Recommendations for Future Work

For all four challenges, in this thesis, we only deal with the simplest possible cases of the general challenges; it is necessary to extend our analysis to more general cases.

- For the first challenge – related to measurement-related certification of systems – we simply propose an idea, it is still necessary to develop this idea and to come up with the corresponding standards.

- For the second challenge – related to measurement design – we only deal with the simplest case when the data processing algorithms consists of simply adding or averaging the measurement results. It is necessary to extend our analysis to more complex data processing algorithms.

- For the third challenge – of uncertainty analysis in situations when we have different information about different pairs of measurements – we only deal with the cases when for the most pairs, we have information of the same type, and only for a small number of pairs, we have different information. It is necessary to extend our analysis to situations when we have a larger number of pairs with different information.

- Finally, for the fourth challenge – related to processing measurement results – we only deal with the case when we know the ordering of the measurement results, but not the numerical values themselves. It is necessary to extend our analysis to situations when we have (and can use) numerical values as well.

References

[1] J. Beirlant, Y. Goegevuer, J. Teugels, and J. Segers, *Statistics of Extremes: Theory and Applications*, Wiley, Chichester, 2004.

[2] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*, Springer Verlag, Berlin, Heidelberg, New York, 2006.

[3] A. Dean, D. Voss, and D. Dragulić, *Design and Analysis of Experiments*, Springer, Cham, Switzerland, 2017.

[4] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer Verlag, Berlin, Heidelberg, New York, 2012.

[5] A. S. Gerber and D. P. Green, *Field Experiments: Design, Analysis, and Interpretation*, W. W. Norton, New York, 2012.

[6] P. Goos and B. Jones, *Optimal Design of Experiments: A Case Study Approach*, Wiley, Chichester, West Sussex, UK, 2011.

[7] E. J. Gumbel, *Statistics of Extremes*, Dover Publ., New York, 2004.

[8] B. Gunter and D. Coleman, *A DOE Handbook:: A Simple Approach to Basic Statistical Design of Experiments*, CreateSpace Independent Publishing Platform, 2014.

[9] ITER Project, https://www.iter.org/

[10] L. Jaulin, M. Kiefer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2001.

[11] B. Jones and D. C. Montgomery, *Design of Experiments: A Modern Approach*, Wiley, Hoboken, New Jersey, 2020.

[12] D. Kramer, "Further delays at ITER are certain, but their duration isn't clear", *Physics Today*, May 2022, pp. 20–22.

[13] B. J. Kubica, *Interval Methods for Solving Nonlinear Contraint Satisfaction, Optimization, and Similar Problems: from Inequalities Systems to Game Solutions*, Springer, Cham, Switzerland, 2019.

[14] J. Lawson, *Design and Analysis of Experiments with R*, Chapman & Hall/CRC, Boca Raton, Florida, 2015.

[15] J. Lorkowski, O. Kosheleva, V. Kreinovich, and S. Soloviev, "How design quality improves with increasing computational abilities: general formulas and case study of aircraft fuel efficiency", *Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII)*, 2015, Vol. 19, No. 5, pp. 581–584.

[16] G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.

[17] J. Momoh, *Smart Grid: Fundamentals of Design and Analysis*, IEEE Press, Piscataway, New Jersey, 2012.

[18] D. C. Montgomery, *Design and Analysis of Experiments*, Wiley, Hoboken, New Jersey, 2020.

[19] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.

[20] S. G. Rabinovich, *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, New York, 2005.

[21] K. Rekab and M. Shaikh, *Statistical Design of Experiments with Engineering Applications*, CRC Press, Boca Raton, Florida, 2019.

[22] A. Roben, *Handbook of Design and Analysis of Experiments*, Chapman & Hall/CRC, Boca Raton, Florida, 2012.

[23] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

Curriculum Vita

Hector Alejandro Reyes received his Bachelor in Computer Science degree from UTEP in Fall 2021. In Spring 2021, he was Instructional Assistant for UTEP's Department of Computer Science. From Summer 2021 until Summer 2022, he was Research Assistant for UTEP's Department of Industrial, Manufacturing, and Systems Engineering.

During his studies at UTEP, he has co-authored two papers that were accepted:

H. Reyes, D. Trinh, and V. Kreinovich, "Fault detection in a smart electric grid: geometric analysis", In: M. Ceberio and V. Kreinovich (eds.), *Decision Making under Uncertainty and Constraints: A Why-Book*, Springer, Cham, Switzerland, 2022, to appear.

H. Reyes, S. Tizpaz-Niari, and V. Kreinovich, "Over-measurement paradox: suspension of thermonuclear research center and need to update standards", in: M. Ceberio and V. Kreinovich (eds.), *Uncertainty, Constraints, and Decision Making*, Springer, Cham, Switzerland, 2023, to appear.

He is also finalizing several other research papers that will be submitted for publication soon.