

Die approbierte Originalversion dieser Dissertation ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (<http://www.ub.tuwien.ac.at>).

The approved original version of this thesis is available at the main library of the Vienna University of Technology (<http://www.ub.tuwien.ac.at/englweb/>).

Diese Dissertation haben begutachtet

---



TECHNISCHE  
UNIVERSITÄT  
WIEN  
Vienna University of Technology

DISSERTATION

# Mathematical Modeling for New Insights into Epidemics by Herd Immunity and Serotype Shift

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines  
Doktors der technischen Wissenschaften unter der Leitung von

ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Felix Breitenecker

E101 - Institut für Analysis und Scientific Computing

eingereicht an der Technischen Universität Wien  
Fakultät für Mathematik und Geoinformation

von

Dipl.-Ing. Florian Matthias Miksch

Matrikelnummer: 0326203  
Schröttergasse 3/14, 1100 Wien, Österreich

Wien, im Oktober 2012



# Kurzfassung

**Motivation:** Die Simulation und Berechnung von Epidemien hat eine lange Tradition in der Mathematik, aber auch in anderen Wissenschaften. Dominierten früher noch einfache Laborversuche und Modelle, werden die Simulationen durch steigende Rechenleistungen der Computer und Verfügbarkeit von umfassenderen Daten zunehmend komplexer und detaillierter. Dadurch können Fragestellungen untersucht werden, deren Betrachtung bisher nicht möglich war.

**Ziele und Methoden:** Das Ziel dieser Arbeit ist die Untersuchung, Definition und Interpretation von dynamischen Effekten bei der Ausbreitung von Epidemien unter Verwendung verschiedener Modelle. Wesentlich ist dabei der gezielte Einsatz und Vergleich von agentenbasierten Modellen, welche die Ausbreitung der Krankheit als Folge von Übertragungen von Mensch zu Mensch simulieren und Markov-Modellen, die auf stochastischen Prozessen basieren. Während erstere durch ihre Struktur in der Lage sind, reale Effekte zu erzeugen, verwenden letztere vorgegebene Übergangswahrscheinlichkeiten und erlauben daher keine Rückkopplungs-Effekte.

**Ergebnisse:** Einfache Epidemien bestehen aus einem einzelnen Erreger, welcher sich in einer Bevölkerung ausbreitet. In komplexeren Fällen treten mehrere Arten eines Erregers, sogenannte Serotypen, auf, die miteinander in Konkurrenz stehen und sich gegenseitig in ihrer Ausbreitung behindern. Interventionen gegen derartige Epidemien, beispielsweise durch Impfungen, haben im Allgemeinen nichtlineare Effekte zur Folge. Dieses nichtlineare Systemverhalten wurde in der Realität in der Vergangenheit regelmäßig beobachtet und dokumentiert und tritt auch im agentenbasierten Modell auf. Ein wichtiger Effekt in diesem Zusammenhang ist die Herdenimmunität, die den Vorteil einer Impfstrategie für ungeimpfte Personen beschreibt. Im Fall konkurrierender Serotypen kann es durch Impfungen außerdem zu einer Änderung der Kräfteverhältnisse der Serotypen kommen, welche als Serotypenverschiebung bezeichnet wird. Durch geeignete, flexible Definitionen sind aussagekräftige Interpretationen dieser Phänomene möglich, die zusätzliche Erkenntnisse über die Ausbreitung von Epidemien liefern. Markov-Modelle sind aufgrund vorgegebener Übergangswahrscheinlichkeiten nicht in der Lage, derartige Effekte zu erzeugen und liefern daher vom agentenbasierten Modell abweichende Ergebnisse.

## KURZFASSUNG

---

Durch Berücksichtigung von Herdenimmunität und, im Fall konkurrierender Serotypen, auch Serotypenverschiebung im Markov-Modell können allerdings die Ergebnisse des agentenbasierten Modells erreicht werden. Dieser Zusammenhang wird anhand einiger Testfälle gezeigt und lässt sich auch analytisch nachweisen.

**Fazit:** Die Definitionen in dieser Arbeit sind hinreichend flexibel, sodass sie sowohl bei Modellen als auch unter Laborbedingungen und bei realen Epidemien verwendbar sind. Es zeigt sich deutlich, dass agentenbasierte und Markov-Modelle unterschiedliche Resultate liefern, welche allerdings erklärt und quantitativ dargestellt werden können. Die Ergebnisse führen zu dem Schluss, dass Herdenimmunität und Serotypenverschiebung die Gesamtheit aller dynamischen Effekte darstellen, die durch Interventionen gegen Epidemien verursacht werden.

Durch die entwickelten Methoden und gewonnenen Erkenntnisse werden die Auswirkungen von Impfstrategien besser verständlich und können potentiell genauer prognostiziert werden. Darüber hinaus liefern sie Hinweise, wie ein Modell beschaffen sein muss, um eine konkrete Epidemie bestmöglich zu simulieren und um Impfungen und andere Maßnahmen korrekt und valide abschätzen zu können. Agentenbasierte Modelle sind empfehlenswert aufgrund ihrer Modellierung von realen Strukturen und ihrer Fähigkeit, entstehendes Verhalten darzustellen. Markov-Modelle dagegen sind nur bedingt und in wenigen Situationen zur Simulation ansteckender Krankheiten geeignet.

# Abstract

**Motivation:** Models for simulation of epidemics have a long tradition in the field of mathematics and other sciences. In the past, simple laboratory experiments and models were mainly used. But today, simulations are getting more complex and detailed due to increasing calculation power of computers and availability of more comprehensive data. Hence it is possible to consider questions that could not be examined before.

**Goals and methods:** The aim of this work is to examine, define and interpret dynamic effects on the spread of epidemics through the use of different models. This approach is heavily based on the deliberate application of agent based models, which simulate the spread of a disease as a consequence of transmission from person to person, and Markov models, which are based on stochastic processes. Thanks to its structure, the former approach is able to create real effects, while the latter approach uses given transition probabilities, and hence is not able to produce feedback effects.

**Results:** Simple epidemics consist of a single pathogen that spreads among a population. In more complex cases, several types of the pathogen, which are called serotypes, occur. Often they are in competition and interfere with each other in their propagation. Interventions against such epidemics, for example by vaccinations, generally cause nonlinear effects. This nonlinear system behavior has been observed and documented several times in the real world and also occurs in agent based models. One important effect in this context is herd immunity which describes the benefit of unvaccinated people by a vaccination strategy. In case of competitive serotypes, vaccinations might additionally change the balance of powers of serotypes, which is called serotype shift. Suitable, flexible definitions allow meaningful interpretations of these phenomena that provide additional information on the spread of epidemics. Markov models do not allow generating such effects due to given transition probabilities, hence their results differ from the ones of the agent based model. By taking into consideration the herd immunity and, in case of competitive serotypes, also the serotype shift, the results of the Markov model can be corrected so that they match the agent based model. This relationship is demonstrated in representative test cases and is also proved analytically.

## ABSTRACT

---

**Conclusions:** The definitions in this work are sufficiently flexible so that they can be used for models as well as for experiments under laboratory conditions as well as on real epidemics. The differences between agent based and Markov models are obvious but it is possible to explain them clearly and quantitatively. This leads to the conclusion that herd immunity and serotype shift cover all dynamic effects that are caused by interventions against an epidemic.

By the developed methods and gathered insights, the consequences of vaccination strategies become more understandable and can potentially be predicted more accurately. Furthermore, they lead to hints on how a model needs to be set up so that it simulates epidemics best and can be used to assess the impact of vaccinations and other interventions correctly and validly. Agent based models seem to be suitable because of their real world structure and their ability to capture emerging behavior. Markov models, in contrast, are useable only in limited situations and under restrictive assumptions for simulation of infectious diseases.

# Contents

<b>Kurzfassung</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Contents</b>	<b>v</b>
<b>List of figures</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Agent based models</b>	<b>3</b>
2.1 Historical background . . . . .	3
2.2 Goals of agent based modelling . . . . .	3
2.3 Notations . . . . .	4
2.4 What is an agent? . . . . .	4
2.4.1 Conceptual analysis . . . . .	4
2.4.2 Definitions in literature . . . . .	4
2.4.3 Summary . . . . .	6
2.5 Characteristics of agent based models . . . . .	7
2.5.1 Modelling of emerging behavior . . . . .	7
2.5.2 Natural description of a system . . . . .	10
2.5.3 Flexibility . . . . .	10
2.6 Advantages and disadvantages of agent based modeling . . . . .	10

<b>3</b>	<b>Different approaches for modeling of epidemics</b>	<b>13</b>
3.1	Introduction . . . . .	13
3.2	Terms . . . . .	14
3.3	The basic model types: SIS and SIR . . . . .	14
3.3.1	SIS models . . . . .	14
3.3.2	SIR models . . . . .	15
3.3.3	Extensions of SIS and SIR models . . . . .	16
3.4	Top-Down & Bottom-Up . . . . .	17
3.5	Markov models . . . . .	17
3.5.1	Theoretical basics . . . . .	18
3.5.2	From the theory to the model . . . . .	20
3.5.3	Discussion . . . . .	24
3.5.4	SIS epidemics with Markov models . . . . .	26
3.5.5	SIR models with Markov models . . . . .	26
3.6	Differential equations . . . . .	27
3.6.1	SIS models with differential equations . . . . .	27
3.6.2	SIR models with differential equations . . . . .	29
3.7	Cellular Automata . . . . .	30
3.8	Agent based models . . . . .	32
3.8.1	Example of an agent based epidemic model . . . . .	32
3.8.2	Extension of agent based epidemic models . . . . .	34
3.9	Comparison of the presented approaches . . . . .	35
3.9.1	Markov models . . . . .	35
3.9.2	Differential equations . . . . .	36
3.9.3	Cellular automata . . . . .	36
3.9.4	Agent based models . . . . .	37
<b>4</b>	<b>Herd immunity: Terminology, motivation and origin</b>	<b>39</b>
4.1	Usage of the term “herd immunity” in the literature . . . . .	39



---

4.2	Usage of the term “herd immunity” in this work . . . . .	40
4.3	Emergence of herd immunity . . . . .	42
4.4	Calculation of the herd immunity . . . . .	43
4.4.1	Definition . . . . .	43
4.4.2	Quantification of the disadvantage . . . . .	45
4.4.3	Alternative calculation of the disadvantage . . . . .	46
4.5	Herd immunity in models . . . . .	47
4.5.1	Usage . . . . .	47
4.5.2	Prerequisites . . . . .	47
4.6	Herd immunity as an explaining quantity . . . . .	48
4.6.1	Introduction . . . . .	48
4.6.2	Analytic calculation of herd immunity in an SIS model . . . . .	49
4.6.3	Example: Herd immunity in an agent based SIS model . . . . .	52
4.6.4	Example: Herd immunity in an agent based SIR model . . . . .	57
4.7	Conclusions . . . . .	58
4.8	Parameter studies on herd immunity . . . . .	60
4.8.1	Relationships in the SIS model . . . . .	60
4.8.2	Parameter studies with the SIS model . . . . .	63
4.8.3	Parameter studies with the SIR model . . . . .	69
<b>5</b>	<b>Comparison of Markov models and agent based models for epidemic simulations</b>	<b>73</b>
5.1	Overview . . . . .	73
5.2	Comparison of equilibriums in an SIS epidemic . . . . .	74
5.2.1	Construction of the models . . . . .	74
5.2.2	Parameterization of the model . . . . .	75
5.2.3	Implementation of vaccinations . . . . .	76
5.2.4	Explanation of the differences and harmonization of the results . . .	77
5.3	Comparison of the temporal progress of an SIR epidemic . . . . .	78
5.3.1	Construction of the models . . . . .	78

## CONTENTS

---

5.3.2	Parameterization of the model . . . . .	79
5.3.3	Implementation of vaccinations . . . . .	81
5.3.4	Explanation of the differences and harmonization of the results . . .	82
5.4	Comparison of a complete SIR epidemic . . . . .	84
5.4.1	Construction of the models . . . . .	84
5.4.2	Parameterization of the model . . . . .	84
5.4.3	Implementation of vaccinations . . . . .	84
5.4.4	Explanation of the differences and harmonization of the results . . .	85
5.5	Analytic considerations . . . . .	86
5.6	Conclusions . . . . .	86
<b>6</b>	<b>Simulation of two competitive serotypes</b>	<b>89</b>
6.1	Consideration of serotypes . . . . .	89
6.2	Competition of serotypes . . . . .	90
6.2.1	Streptococcus pneumoniae . . . . .	90
6.2.2	Influenza . . . . .	90
6.2.3	Haemophilus Influenzae . . . . .	91
6.2.4	Serotype shift and consequences of competitive serotypes . . . . .	91
6.3	Models with competitive serotypes in literature . . . . .	91
6.4	Preliminary remarks on modeling of competitive serotypes . . . . .	93
6.5	Modeling approach 1: Competition by carrying only one serotype . . . . .	94
6.5.1	Modeling an SIS epidemic with approach 1 . . . . .	95
6.5.2	Modeling an SIR epidemic with approach 1 . . . . .	97
6.6	Modeling approach 2: Competition by carrying only one serotype and mutual susceptibility . . . . .	100
6.6.1	Modeling an SIS epidemic with approach 2 . . . . .	100
6.6.2	Modeling an SIR epidemic with approach 2 . . . . .	103
6.6.3	Conclusions of modeling approach 2 . . . . .	103

---

<b>7</b>	<b>Herd immunity and serotype shift in case of two competitive serotypes</b>	<b>105</b>
7.1	Initial situation . . . . .	105
7.2	Requirements and preliminary remarks . . . . .	105
7.3	Definitions of herd immunity and serotype shift . . . . .	107
7.4	Definition of the disadvantage . . . . .	110
7.5	Herd immunity and serotype shift as an explaining quantity . . . . .	110
7.5.1	Herd immunity and serotype shift in SIS models . . . . .	111
7.5.2	Herd immunity and serotype shift in SIR models . . . . .	112
7.6	Parameter studies for herd immunity and serotype shift . . . . .	117
7.6.1	Parameter variations of the SIS model . . . . .	117
7.6.2	Parameter variations of the SIR model . . . . .	118
7.7	Conclusions . . . . .	122
<b>8</b>	<b>Comparison of Markov and agent based models for epidemic simulations with two competitive serotypes</b>	<b>127</b>
8.1	Overview . . . . .	127
8.1.1	Introduction . . . . .	127
8.1.2	Considerations about comparability of agent based and Markov models with two serotypes . . . . .	128
8.1.3	The models . . . . .	128
8.2	Comparison of equilibriums in an SIS epidemic . . . . .	129
8.2.1	Construction of the models . . . . .	129
8.2.2	Parameterization of the models . . . . .	130
8.2.3	Implementation of vaccinations . . . . .	133
8.2.4	Explanation of the differences and harmonization of the results . . .	133
8.3	Comparison of the temporal progress of an SIR epidemic . . . . .	133
8.3.1	Construction of the models . . . . .	133
8.3.2	Parameterization of the models . . . . .	138
8.3.3	Implementation of vaccinations . . . . .	140
8.3.4	Explanation of the differences and harmonization of the results . . .	141

## CONTENTS

---

8.4	Comparison of a whole SIR epidemic . . . . .	142
8.4.1	Construction of the models . . . . .	142
8.4.2	Parameterization of the models . . . . .	143
8.4.3	Implementation of vaccinations . . . . .	144
8.4.4	Explanation of the differences and harmonization of the results . . .	144
8.5	Analytic considerations . . . . .	146
8.6	Conclusions . . . . .	147
<b>9</b>	<b>Summary and Conclusions</b>	<b>149</b>
9.1	Summary of insights . . . . .	149
9.2	Conclusions . . . . .	152
9.3	Further considerations . . . . .	153
9.4	Recommendation for usage of Markov models . . . . .	154
	<b>Bibliography</b>	<b>155</b>
	<b>Lebenslauf</b>	<b>161</b>

# List of Figures

2.1	Visualization of an agent according to the definition presented at the Wintersim 2005/2006 [38] . . . . .	6
2.2	“emerging behavior”: Rule for example 1 . . . . .	8
2.3	“emerging behavior”: System with rules of example 1 . . . . .	8
2.4	“emerging behavior”: Rule for example 2 . . . . .	9
2.5	“emerging behavior”: System with rules of example 2 . . . . .	9
3.1	SIS model . . . . .	15
3.2	Typical results of SIS models . . . . .	15
3.3	SIR model . . . . .	16
3.4	Typical results of an SIR model . . . . .	17
3.5	Example of a state transition diagram . . . . .	19
3.6	Example of a Markov tree . . . . .	22
3.7	Example of a decision tree . . . . .	23
3.8	Commonly used terms for models in the area of stochastic processes . . . . .	24
3.9	Result of a typical SIS Markov model . . . . .	26
3.10	Result of a typical SIR Markov model . . . . .	27
3.11	Visualization of the differential equation SIS model . . . . .	28
3.12	Typical result of a differential equation SIS model . . . . .	29
3.13	Visualization of the differential equation SIR model . . . . .	29
3.14	A typical result of an SIS model of differential equations . . . . .	30
3.15	Neighborhoods on cellular automata . . . . .	31
3.16	Realization of an SIR model as a cellular automaton . . . . .	31

## LIST OF FIGURES

---

3.17	Realization of an SIR model based on an LGCA . . . . .	32
3.18	Agents in an epidemic model . . . . .	33
3.19	Contacts between agents in an epidemic model . . . . .	33
3.20	Infections of agents in an epidemic model based on contacts . . . . .	33
3.21	Recovery of agents in an epidemic model . . . . .	34
3.22	Result of a typical agent based SIS model . . . . .	34
3.23	Result of a typical agent based SIR model . . . . .	34
4.1	Visualization of the infection probability of a marked person . . . . .	42
4.2	Infection states of persons in an SIS model . . . . .	48
4.3	Infection states of persons in an SIR model . . . . .	49
4.4	Graph of the herd immunity depending on the vaccinated part of the population . . . . .	52
4.5	States in the SIS model . . . . .	53
4.6	New infections in the SIS model . . . . .	55
4.7	Calculation of the herd immunity upon current new infections . . . . .	55
4.8	Calculation of the herd immunity upon infections since simulation start . . . . .	55
4.9	Calculation of the herd immunity upon current carriers . . . . .	55
4.10	States in the SIS model with vaccinations after 100 time steps . . . . .	56
4.11	Infections in the SIS model with vaccinations after 100 time steps . . . . .	56
4.12	Herd immunity based on current infections with vaccinations after 100 time steps . . . . .	56
4.13	States in the SIR model . . . . .	57
4.14	New infections in the SIR model . . . . .	59
4.15	Is there a disadvantage or still a benefit in time 40 to 80? . . . . .	59
4.16	Calculation of the herd immunity upon current new infections . . . . .	59
4.17	Calculation of the herd immunity upon infections since simulation start . . . . .	59
4.18	Calculation of the herd immunity upon current carriers . . . . .	59
4.19	Results of the test series for infection probability and contacts per persons. . . . .	61
4.20	Deviation of the agent based model from the analytic calculation in case of no vaccinations. . . . .	62

4.21	Deviation of the agent based model from the analytic calculation in case of 10% vaccination rate. . . . .	62
4.22	Deviation of the agent based model from the analytic calculation in case of 20% vaccination rate. . . . .	62
4.23	Deviation of the agent based model from the analytic calculation in case of 30% vaccination rate. . . . .	62
4.24	Variation of the infection probability in the SIS model . . . . .	65
4.25	Herd immunity of the variation of the infection probability in the SIS model	65
4.26	Variation of the recovery time in the SIS model . . . . .	65
4.27	Herd immunity of the variation of the recovery time in the SIS model . . .	66
4.28	Variation of the infection probability in the SIR model . . . . .	70
4.29	Herd immunity of the variation of the infection probability in the SIR model	70
4.30	Variation of the recovery time in the SIR model . . . . .	71
4.31	Herd immunity of the variation of the recovery time in the SIR model . . .	71
5.1	The state transition diagram of the SIS Markov model for unvaccinated people . . . . .	74
5.2	The Markov tree of the SIS Markov model . . . . .	75
5.3	Carrier rates of the equally parameterized SIS models without vaccinations	76
5.4	Results of vaccinations in the Markov model and in the agent based model for SIS epidemics . . . . .	77
5.5	Results of vaccinations in the Markov model (corrected by the herd immunity) and in the agent based model for SIS epidemics . . . . .	78
5.6	The state transition diagram of the SIR Markov model for unvaccinated people . . . . .	79
5.7	The decision tree of the SIR Markov model . . . . .	80
5.8	Time-dependent parameter in the SIR Markov model without vaccination .	81
5.9	Results of the equally parameterized SIR models . . . . .	81
5.10	Results of vaccinations in the Markov model and in the agent based model for SIR epidemics . . . . .	82
5.11	Herd immunity in the agent based SIR model based on the current carrier rate . . . . .	83
5.12	Results of vaccinations in the Markov model (corrected by the herd immunity) and in the agent based model for SIR epidemics . . . . .	83

## LIST OF FIGURES

---

5.13	Relationship between agent based models and Markov models through herd immunity . . . . .	87
6.1	Visualization of an SIS disease with two competitive serotypes (using approach 1) . . . . .	95
6.2	Carrier rates of an SIS epidemic with equivalent serotypes and approach 1 ( $p = 0.02, d = 10$ time steps) . . . . .	96
6.3	Carrier rates of an SIS epidemic with different serotypes and approach 1 ( $p_1 = 0.022, p_2 = 0.02, d = 10$ time steps) . . . . .	96
6.4	Carrier rates of an SIS epidemic with different serotypes and approach 1 ( $p = 0.02, d_1 = 11$ time steps, $d_2 = 10$ time steps) . . . . .	96
6.5	Visualization of an SIR disease with competitive serotypes (using approach 1) . . . . .	98
6.6	Carrier rates of an SIR epidemic with equivalent serotypes and approach 1 ( $p = 0.02, d = 10$ time steps) . . . . .	99
6.7	Carrier rates of an SIR epidemic with different serotypes and approach 1 ( $p_1 = 0.022, p_2 = 0.02, d = 10$ time steps) . . . . .	99
6.8	Carrier rates of an SIR epidemic with different serotypes and approach 1 ( $p = 0.02, d_1 = 11$ time steps, $d_2 = 10$ time steps) . . . . .	99
6.9	Visualization of an SIS disease with two competitive serotypes (using approach 2) . . . . .	100
6.10	Visualization of an SIR disease with two competitive serotypes (using approach 2) . . . . .	101
6.11	Carrier rates of an SIS epidemic with equivalent serotypes and approach 2 ( $p = 0.02, d = 10$ time steps) . . . . .	102
6.12	Carrier rates of an SIS epidemic with different serotypes and approach 2 ( $p_1 = 0.022, p_2 = 0.02, d = 10$ time steps) . . . . .	102
6.13	Carrier rates of an SIS epidemic with different serotypes and approach 2 ( $p = 0.02, d_1 = 11$ time steps, $d_2 = 10$ time steps) . . . . .	102
6.14	Carrier rates of an SIR epidemic with equivalent serotypes and approach 2 ( $p = 0.02, d = 10$ time steps) . . . . .	103
6.15	Carrier rates of an SIR epidemic with different serotypes and approach 2 ( $p_1 = 0.022, p_2 = 0.02, d = 10$ time steps) . . . . .	103
6.16	Carrier rates of an SIR epidemic with different serotypes and approach 2 ( $p = 0.02, d_1 = 11$ time steps, $d_2 = 10$ time steps) . . . . .	103
7.1	Visualizing the emergence of serotype shift . . . . .	107



7.2	New infections by serotype in the SIS simulations . . . . .	111
7.3	Total new infections in the SIS simulations . . . . .	111
7.4	Carrier rates by serotype in the SIS simulations . . . . .	112
7.5	Total carrier rates in the SIS simulations . . . . .	112
7.6	Epidemic effects in the SIS simulation, calculated upon current new infections	113
7.7	Epidemic effects in the SIS simulation, calculated upon new infections since start . . . . .	113
7.8	Epidemic effects in the SIS simulation, calculated upon the carrier rate . .	113
7.9	New infections in the SIS simulation with vaccination after 100 time steps .	114
7.10	Carrier rates in the SIS simulation with vaccination after 100 time steps . .	114
7.11	Epidemic effects in the SIS simulation with vaccination after 100 time steps, calculated upon current new infections . . . . .	114
7.12	New infections by serotype in the SIR simulation . . . . .	115
7.13	Total new infections in the SIR simulation . . . . .	115
7.14	Carrier rates by serotype in the SIR simulation . . . . .	115
7.15	Total carrier rates in the SIR simulation . . . . .	115
7.16	Epidemic effects in the SIR simulation, calculated upon current new infections	116
7.17	Epidemic effects in the SIR simulation, calculated upon new infections since start . . . . .	116
7.18	Epidemic effects in the SIR simulation, calculated upon the carrier rate . .	116
7.19	New infections per time step (variation of the infection probability in the SIS model) . . . . .	118
7.20	Carrier rates (variation of the infection probability in the SIS model) . . .	119
7.21	The herd immunity (variation of the infection probability in the SIS model)	119
7.22	The serotype shift (variation of the infection probability in the SIS model)	119
7.23	The general herd immunity (variation of the infection probability in the SIS model) . . . . .	120
7.24	New infections per time step (variation of the recovery time in the SIS model)	120
7.25	Carrier rates (variation of the recovery time in the SIS model) . . . . .	120
7.26	The herd immunity (variation of the recovery time in the SIS model) . . .	121
7.27	The serotype shift (variation of the recovery time in the SIS model) . . . .	121

## LIST OF FIGURES

---

7.28	The general herd immunity (variation of the recovery time in the SIS model)	121
7.29	Total number of infections (variation of the infection probability in the SIR model) . . . . .	123
7.30	The herd immunity (variation of the infection probability in the SIR model)	123
7.31	The serotype shift (variation of the infection probability in the SIR model)	123
7.32	The general herd immunity (variation of the infection probability in the SIR model) . . . . .	124
7.33	Total number of infections (variation of the recovery time in the SIR model)	124
7.34	The herd immunity (variation of the recovery time in the SIR model) . . .	124
7.35	The serotype shift (variation of the recovery time in the SIR model) . . . .	125
7.36	The general herd immunity (variation of the recovery time in the SIR model)	125
8.1	The state transition diagram of a Markov model in case of two competitive serotypes . . . . .	130
8.2	Tree structure of a Markov model in case of two competitive serotypes . . .	131
8.3	Carrier rates of the agent based model compared to the Markov model in the equilibrium of the SIS simulation . . . . .	134
8.4	Herd immunity and serotype shift in the equilibrium of the SIS model . . .	134
8.5	Carrier rates of the agent based model compared to the adjusted Markov model in the equilibrium of the SIS simulation . . . . .	134
8.6	The state transition diagram of the Markov model for SIR epidemics with two competitive serotypes . . . . .	137
8.7	Transition matrix of the Markov model for SIR epidemics with two competitive serotypes for unvaccinated people . . . . .	139
8.8	Transition matrix of the Markov model for SIR epidemics with two competitive serotypes for vaccinated people . . . . .	139
8.9	The carrier rates of the agent based and the Markov model for SIR epidemics in the base case without vaccinations . . . . .	140
8.10	The carrier rates of the agent based and the Markov model for SIR epidemics in the test case with 10% vaccinations . . . . .	140
8.11	The carrier rates of the agent based and the Markov model for SIR epidemics in the test case with 20% vaccinations . . . . .	140
8.12	The carrier rates of the agent based and the Markov model for SIR epidemics in the test case with 30% vaccinations . . . . .	140

---

8.13	Herd immunity in the SIR model . . . . .	141
8.14	Serotype shift in the SIR model . . . . .	141
8.15	The carrier rates of the agent based and the corrected Markov model for SIR epidemics in the test case with 10% vaccinations . . . . .	141
8.16	The carrier rates of the agent based and the corrected Markov model for SIR epidemics in the test case with 20% vaccinations . . . . .	142
8.17	The carrier rates of the agent based and the corrected Markov model for SIR epidemics in the test case with 30% vaccinations . . . . .	142
8.18	State transition diagram of the Markov Model of a whole SIR season . . .	143
8.19	Relationship of vaccinations between agent based models and Markov mod- els with two serotypes. . . . .	148
9.1	Stochastic and contact based epidemic models in comparison . . . . .	152
9.2	Stochastic and contact based epidemic models in comparison . . . . .	153

## LIST OF FIGURES

---

# Chapter 1

## Introduction

Models for simulation of epidemics have a long tradition. Humans have always been interested in infectious diseases because they have consistently taken their toll on human life and they are still doing it. In addition to that, they cause numerous sicknesses with partly severe consequences. In general, one might talk about an epidemic when the disease affects many people. A better definition does not exist and actually there is no need to have one. The term “epidemic” arises from a necessity based on different situations and not least it heavily depends on the public awareness.

**The analysis of epidemics pursues at least two goals:**

- To know, why and how an epidemic is spread
- To come up with the most realistic estimates of consequences of interventions to fight epidemics accurately and efficiently

First mathematical approaches of simulation of infectious diseases hark back to the 1920s when A. G. McKendrick and W. O. Kermack implemented their SIR models with differential equations [28]. Today, 80 years later, these models are still popular and widely used. At the same time, other scientists performed laboratory experiments with groups of animals to examine the spread of a disease under standardized conditions [53]. In contrast to statistical calculations, the experiments revealed effects on the epidemics that could not be expected a priori. The dynamics of epidemics also occurred in the real world. While smallpox could be eradicated and measles could be significantly reduced with comprehensive vaccination strategies, another ambitious program to fight the Malaria failed. Despite the knowledge about dynamic effects on epidemics, Markov Models became the most important tool to predict occurrence of epidemics and to examine strategies. Although they are not able to cover such effects, they have been used due to their growing popularity in the past few decades in the area of “evidence based medicine”. At the end of the 1990s, computational power grew rapidly and simulation of new and more complex models became possible. These approaches simulate single persons who live in an environment and can infect each other. Such models do not rely on knowledge about the

epidemic itself but instead on its fundamental basics. These are the transmissions based upon contacts and the progression of the disease. The epidemic then occurs automatically. These models offer totally new opportunities to examine the spread of epidemics. Moreover, it is possible to analyze effects that are suspected to occur in reality but could not be examined properly under standardized conditions, even if they might have been observed in special cases.

This work gives special attention to two such effects: herd immunity and serotype shift. Herd immunity refers to the phenomenon in case of vaccinations against a disease. Not protected people meet less infected people due to vaccinated people in the population, consequently they have a lower risk of getting infected and so they profit from the vaccinations as well. These consequences eventually result in a different propagation of the epidemic. These connections can even lead to a situation where an epidemic cannot spread any more even though only a part of the population is vaccinated. Serotype shift becomes relevant for diseases that are caused by different variants of one pathogen just like in the case of influenza or streptococcus pneumoniae. These variants, often called serotypes, compete for susceptible persons and interfere each other as they spread. Vaccines often work against some serotypes only. Such an intervention in a population result in a larger living environment for the other serotypes, hence they can spread more freely. This can reduce the benefit of a vaccination strategy, even though the vaccines work and fight the designated serotypes successfully. Both phenomena are heavily disputed and their real occurrence is widely unknown. Relevant literature does not even provide concrete definitions which could put the discussion on a sound basis.

The main goal of this work is to examine herd immunity and serotype shift. First of all, it is necessary to find general, useful and mathematical definition of these effects. Second, the task is to observe the degree of these phenomena in agent based models which are close to reality due to their contact based epidemic spread. Furthermore, one might suppose that herd immunity occurs for the reason of contact based transmission. This is precisely what is modeled by the agents. Previous studies show definitely that both herd immunity and serotype shift occur in agent based models. Analysis of test cases lead to an assessment of practicability of the definitions and also to extensive and valid interpretations due to quantitative results. Situations where herd immunity become so strong that proper spread is not possible any more and the disease gets extinct will be considered particularly. Another task is to compare contact based and stochastic models by using agent based and Markov models. Stochastic models always rely on transmission probabilities which must be explicitly known and completely describe the behavior of the whole system. This comparative study is pursuing two goals. For one thing it should reveal the structural differences of those two model types and outline differences in the results. For another thing, it should try to find out whether it is possible to consider herd immunity and serotype shift in Markov models and how this could be accomplished. Firstly, this leads to a deeper understanding on phenomena in the spread of epidemics. Secondly, it can show when and under what circumstances the usage of stochastic models is possible, useful and valid.

# Chapter 2

## Agent based models

Agent based modelling is a relatively young discipline that has become possible with powerful computers in the nineties of the twentieth century.

### 2.1 Historical background

The historical background of agent based models and simulation lies in the “complex adaptive systems” (=CAS). They are so called “bottom up” systems which, in contrast to “top down” systems, are described by the behavior of single elements. The usual background of CAS is the question how complex system behavior emerges and works through connection and interaction of single elements. This includes, among others, self-organization and dynamic reorganization of single components [38].

Relevant in this context is the attempt to understand and simulate people with their individual, social behavior based on independent decisions.

### 2.2 Goals of agent based modelling

- Generation of an environment which allows to model details and relations satisfyingly exact.
- Creation of dynamic effects that cannot be created with other models and that cannot be represented by parameters.
- Build models with natural and descriptive structure and data (in contrast to many other model types that use abstract and aggregated variables).

## 2.3 Notations

Literature shows a variety of different notations when someone talks about agent based models. Common designations are:

- ABM (agent-based modelling)
- ABS (agent-based systems)
- IBM (individual-based modelling)
- MAS (multi-agent systems)
- ABMS (agent-based modelling and simulation)

## 2.4 What is an agent?

### 2.4.1 Conceptual analysis

One might think about secret agents, detectives or the FBI in the context of “agents”. Agents in models are definitely different things, even though their meanings are based on the same origin.

The word “agent” originates from the Latin word “agere” which means “someone who acts”. In English it is also often used as “someone who acts for another”.

This means that an agent is an entity that does something more or less independently, based on its own rules, sometimes on behalf of another. We see that agents are not necessarily persons, especially in the area of modeling.

### 2.4.2 Definitions in literature

Back in the middle of the 1990s, scientists have regularly worked on general, formal definitions for agents in the area of modeling and simulation. Their explications are partly different, but also have certain ideas in common.

One of the first definitions goes back to **Wooldridge (1997)** [62]. According to him, agents have to have four essential features:

- **Autonomy:** Agents are in a state that is not visible to others. They make decisions based on their state without interventions of human and others.
- **Reactivity:** Agents are situated in an environment, they are able to perceive it and are able to respond to changes that occur in it.



- **Pro-activeness:** Agents are able to act and make decisions that are not triggered by something else. In this case they “take the initiative”.
- **Social ability:** Agents can interact with each other in a specific language that allows them to adapt their decisions.

Another early definition originates from **John L. Casti**, also from 1997. In his book “**Would-Be Worlds: How Simulation is Changing the Frontiers of Science**” [7], he describes two basic requirements for agents that he calls a two-level set of rules:

- On the one hand, agents need rules which describe their action.
- On the other hand, agents also need meta-rules which describe how agents can change their own rules upon they act.

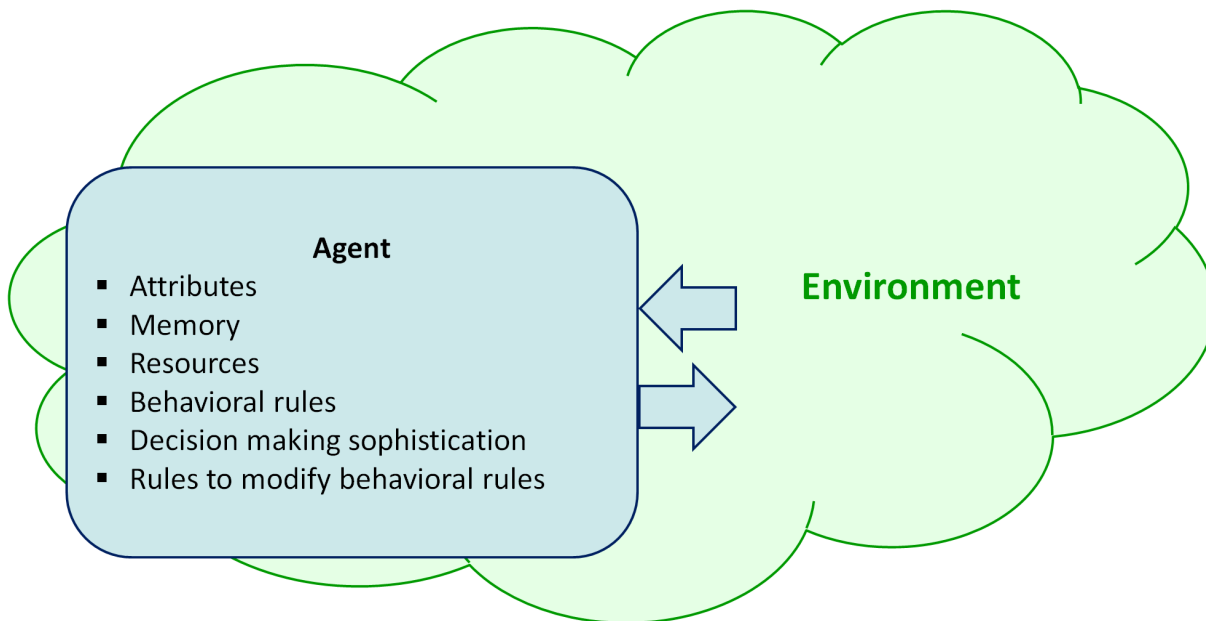
**Jennings (2000)** [24] describes agents generally as components which are able to make independent decisions.

The topic is even more abstract for **Bonabeau (2001)** [4]. In his paper he declares in the broadest sense every model as agent based that consists of components. As an example he takes a system of differential equations which is consequently an agent based model. Each single equation stands for an agent; behavior and interactions of the agents are defined by the right side of the equation.

C. Macal and M. North developed in the context of the **Winter Simulation Conference 2005** a new definition of agents which was revised in 2006 [37] [38]. This work demands an agent to have five features:

- An agent is identifiable.
- An agent lives in an environment with other agents and can interact with both other agents and the environment.
- An agent acts goal-directed.
- An agent acts autonomous and self-directed.
- An agent is flexible, and has the ability to learn and adapt its behaviors over time based on experience.

Figure 2.1 visualizes the technical appearance of agents which also originates from this paper at the Winter Simulation Conference.



**Figure 2.1:** Visualization of an agent according to the definition presented at the Wintersim 2005/2006 [38]

### 2.4.3 Summary

These five attempts of definitions show that it is obviously not easy to clarify how an agent based model should look like in detail.

Why is that?

Agent based models consist, in contrast to classical mathematical models, of a textual, informal description. If the formal description of an agent follows a straight definition, then it could be very easy to change the description slightly and violate the definition. Often, it still makes sense to use such adapted versions of agents. Table 2.1 shows common difficulties in this context.

The questions in table 2.1 illustrate that useful models do not necessarily need to stick to all requirements for agents. The problem is that constructs, that occur by violation of some rules, still look like agents but by definition they are not. Taking account for these considerations, we inevitably end up at Bonabeau's approach which basically allows to see any system as an agent based model as long as it consists of components. Needless to say, it does not make too much sense to call a system of differential equations an agent based model because the mathematical theory is much stronger for that matter. But it leaves open the possibility to model agents that act like differential equations. Such a model can generally be useful and even needed in some cases. Consequently it seems suitable to talk about agents whenever the intention is to work with individual components. These components might be described by at least four features:

- They are clearly identifiable.

Statement	Difficulty
Agents are self-directed.	Can an agent have rules that allows him to be externally controlled? Is not any implementation of agents as a computer program externally controlled?
Agents act goal oriented.	Can the pure obedience of rules already be the goal or must an agent have long-term goals?
Agents have adaptable rules.	Must each and every agent have meta-rules to adapt his behavior?
Agents interact with each other.	Must each and every agent interact with other agents?

**Table 2.1:** Difficulties in the context of definitions for agent based models

- They have properties that might change over time.
- They act upon individual, adaptable rules.
- They can communicate with the environment and other agents.

Additionally, an agent based system requires an appropriate environment in which agents can exist, and with which they can interact.

**Last but not least, this discussion shows clearly that agent based modeling is more a general concept than a specific set of modeling instructions.**

## 2.5 Characteristics of agent based models

The concept of agent based modeling leads to some typical characteristics. Based on Bonabeau [4], they can be differentiated into three categories:

- Modelling of emerging behavior
- Natural description of a system
- Flexibility

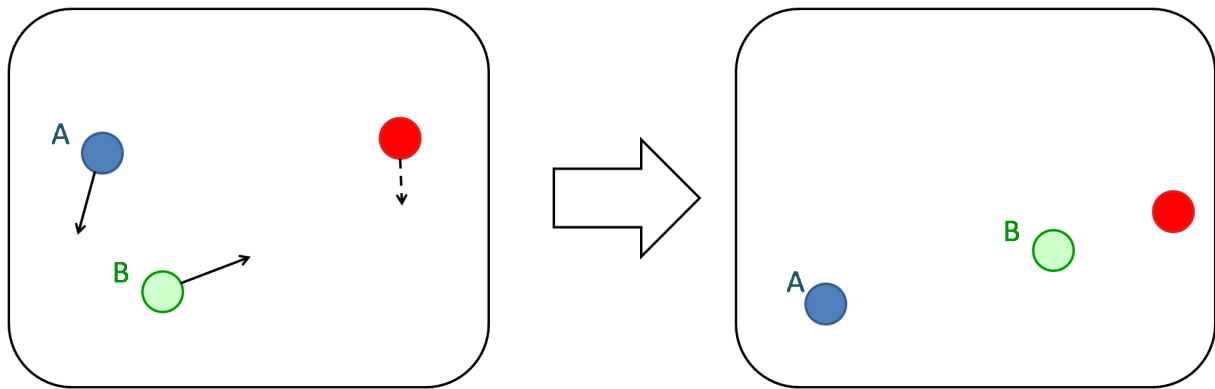
### 2.5.1 Modelling of emerging behavior

An agent based model is described by the behavior of agents. This can lead to a situation where relatively simple rules for agents result in a complex and unpredictable global system behavior. In a way, the system develops its own dynamics. This seems to be a disadvantage but actually, it is a benefit because it allows to model extremely complex systems in which the characteristics are known but the behavior cannot be described appropriately.

**Example 1:** In first simple example we consider a delimited area. The agents are small particles which can move within this area with a certain speed and direction. All agents have only two simple rules:

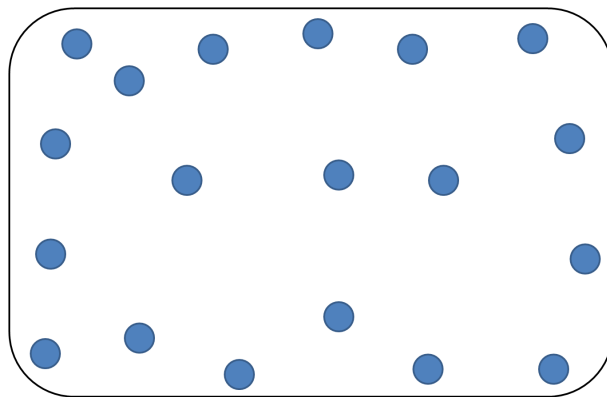
- Choose two random agents A and B at the beginning of the simulation.
- Adapt your speed and direction all the time during the the simulation so that B is always between you and A. Consider position, speed and direction of A, B and yourself.

In other words, the particle should try to be positioned in a line with A and B with B between A and itself. This rule is illustrated in figure 2.2.



**Figure 2.2:** “emerging behavior”: Rule for example 1

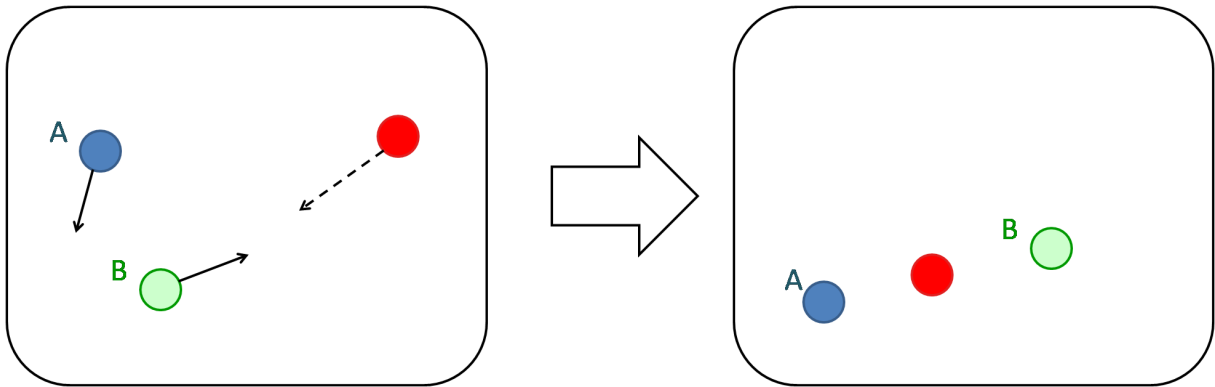
The system behavior of the model is that the particles move around chaotically, mostly close to the border of the area (figure 2.3).



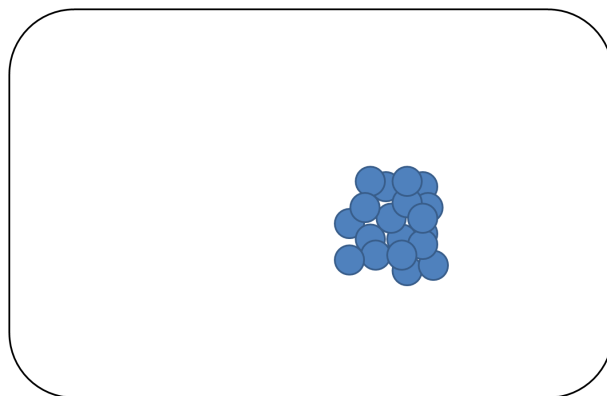
**Figure 2.3:** “emerging behavior”: System with rules of example 1

**Example 2:** Example two is based on example 1 with a different rule 2: The particle has to move, depending on position, speed and direction of A and B, that it is always between A and B (see figure 2.4).

This rule leads to a complete different system behavior: All particles collapse together while this bunch of particles moves around randomly in the area (see figure 2.5).



**Figure 2.4:** “emerging behavior”: Rule for example 2



**Figure 2.5:** “emerging behavior”: System with rules of example 2

## 2.5.2 Natural description of a system

Agent based models provide a natural description of the real system. This means:

- Agents in the model look like components in reality and act in a way that can be observed in the real system.
- Parameters in the model correspond with quantities that occur in the real system.

## 2.5.3 Flexibility

Agent based systems are flexible because of their modular structure. This means many adaptations can be done locally without affecting the rest of the system structure.

The following points are especially important:

- **Scalability:** The number of agents can be increased without limitations while the structure of agents is not affected. Often, an increase of the environment is necessary so that all agents fit in.
- **Level of detail:** The level of detail how agents are modelled is generally not restricted because it depends on the number and complexity of rules and attributes of agents.
- **Rules of single agents** can be adapted easily for the set-up and even during simulation without affecting the structure of other agents and of the environment.
- **Parameterization of single agents** can be set individually and can be changed any time during simulation.

## 2.6 Advantages and disadvantages of agent based modeling

Concluding, we see three main reasons for usage of agent based models:

- **Modeling of real world structures result in real, dynamic behavior.**
- **Easiness of handling and understanding the model.**
- **Structures can be modeled on any detailed level. Results can be examined on any modeled level of details.**

Aside from the fact that the handling and the characteristics of agent based models provide several advantages, there are also some issues that go along with it.

### **Technical advantages:**

- Possibility to model dynamic effects by well-known rules, while the complex dynamic behavior can hardly be described.
- Flexible modelling environment that provides scalability for size and details in a simple way.
- Direct usage of data (parameters, relations between details). Ideally, no adjustment of data is necessary and the whole model gets parameterized in a natural way.
- Extension of agent based models is often less work than extension of other model approaches.

### **Advantages in application:**

- Knowledge about mathematical theories is not necessary.
- Model structure is very clear. Non-modelling project partners (economists, medics, etc.) understand the model structure and can provide important input to the model.

### **Technical disadvantages:**

- Non-consideration of agent rules or wrong agent rules may result in wrong global system behavior. Extensive validation is important to overcome this issue.
- Data for agents (parameters, rules) are sometimes not completely available. Then a comprehensive and often expensive calibration is necessary.
- There are no mathematical methods for analyzing agent-based models. Often, statistical evaluation of results is the only possible method.
- The initial creation of an agent based model can be more programming work than other models.
- Simulations with many agents often require powerful computer systems and long runtimes. Efficient coding and appropriate scaling of the model helps to handle this issue.

### **Dangers in applications:**

- Agent based models look very “real”. The feeling of having a correct model must not replace comprehensive validation.
- Danger of wrong usage and interpretation by non-modelers may happen because the model looks so “real”.





# Chapter 3

## Different approaches for modeling of epidemics

### 3.1 Introduction

There are several methods for modeling and simulation of epidemics. This chapter provides an overview over common modeling techniques and strategies in the area of epidemic simulation. However, according to the context of this work, the presented methods are bound to infectious diseases that are caused by a certain pathogen. Hence, the modeling and interpretation of transmission has a great importance.

On the one hand, this work deals with appropriate simplifications to represent epidemics, particularly with SIS and SIR models.

On the other hand, it describes the most important methods, explains their advantages and disadvantages and demonstrates by examples how epidemics can be modeled with this methods in specific cases.

## 3.2 Terms

susceptible .....	A person that is susceptible to a pathogen.
infected .....	A person that carries a pathogen (=carrier).
infectious .....	A person that is infected with a pathogen and can transmit it to other persons.
resistant .....	A person that cannot get infected with a pathogen.
carrier .....	A person that is “infected”.
sick .....	A person that has symptoms of sickness. Especially for detailed views of diseases, it should be clear that the state “sick” is not necessarily connected to the states “infected” and “infectious”. Thus, it should be used with care.
healthy .....	The opposite of “sick”. Since it is not necessarily connected to the state “susceptible” it should also be used with care.
vaccination rate .....	The vaccination rate indicates the part of the population that is vaccinated or gets vaccinated. In the proper meaning of the word, it is not a rate but only a fraction.
carrier rate .....	The carrier rate indicates the part of infected people in the whole population. Is is also not a rate from a technical point of view.
infection .....	An infection determines the transmission of the disease. Typically, this is the transition from the state “susceptible” to the state “infected”.

## 3.3 The basic model types: SIS and SIR

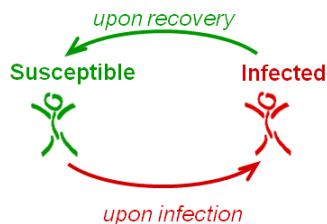
The basic model types SIS and SIR fundamentally have their origins in the epidemic models of W. O. Kermack and A. G. McKendrick which have been published in three papers in the years 1927 to 1933 [28] [29] [30]. Their approaches have been picked up many times and used in numerous variations. They have also been extensively examined and further developed [6, chapter 3].

### 3.3.1 SIS models

SIS models are a simplified representation of a large class of infectious diseases. They represent infections that can be transmitted from person to person. Infected persons recovers after a while but are not able to develop a form of resistance against the disease.

This means that a person is either susceptible (=S) or infected (=I). Susceptible people can get infected upon transmission. Infected people become susceptible again after a while upon recovery (see figure 3.1).

**Characteristics and behavior:** The general characteristic of SIS models is that it ends up in an equilibrium after an adaption phase. This means that the number of susceptible

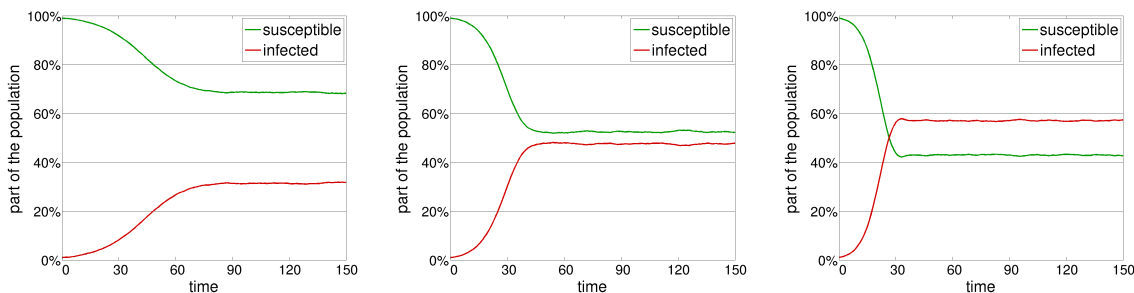


**Figure 3.1:** SIS model

and the number of infected people do not change over time. In general, both numbers are greater than zero. This model state is reached because the number of newly infected people equals the number of recovered people in a certain period of time.

- The epidemic either extincts or reaches a non-trivial equilibrium with infected and susceptible people depending on the parameters.
- The initial values usually do not correspond with the equilibrium. Thus an adaption phase occurs after simulation start and after parameter changes during runtime which ends by reaching the equilibrium. Shape and duration of the adaption phase depend on parameters as well as on the initial state of the model.
- The equilibrium depends on the parameters but it is independent of the initial state. Same parameters always lead to the same equilibrium.
- Tests in previous studies have shown and will show in this work that different modeling approaches of SIS epidemics lead to different results. These differences are based on non-equivalent model structures and parameters that are not completely comparable.

Figure3.2 shows typical simulation results of SIS models.

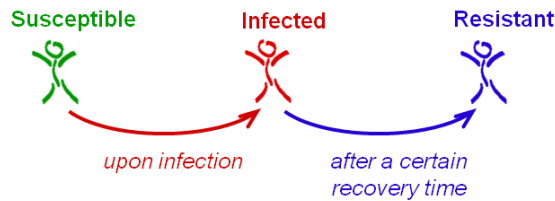


**Figure 3.2:** Typical results of SIS models

### 3.3.2 SIR models

SIR models are, like SIS models, a simplified representation of a large class of infectious diseases. Its only difference from SIS models is that infected persons in an SIR model become resistant against the disease after recovery.

Thus, people are in either of these three available states: susceptible ( $=S$ ), infected ( $=I$ ) and resistant ( $=R$ ). The state  $I$  can only be reached by susceptible persons upon transmission and the state  $R$  can only be reached by infected persons upon recovery (see figure 3.3).



**Figure 3.3:** SIR model

**Characteristics and behavior:** An SIR epidemic always lasts for a limited duration because infected persons become resistant after some time and cannot be infected again. Since the number of persons is finite, there is only a limited pool of susceptible persons. Consequently, only a limited number of persons can be infected with the disease during the simulations and therefore, the epidemic can only last for a finite time. After the epidemic becomes extinct, only susceptible and resistant people are left and transmissions cannot happen any more.

Usually, the initial state of an SIS model is a large susceptible part of the population, a small infected part and no resistant people. This corresponds to the start situation of an epidemic. Typically the number of infected and resistant persons increase rapidly at first, while the number of susceptibles decreases until the number of infected people reaches a peak. Due to the decreasing number of susceptible people after the peak, less people get infected than people who recover. Thus, the number of infected and susceptible people decrease while the resistant people grow. The epidemic is over when only susceptible and resistant but no infected people are left.

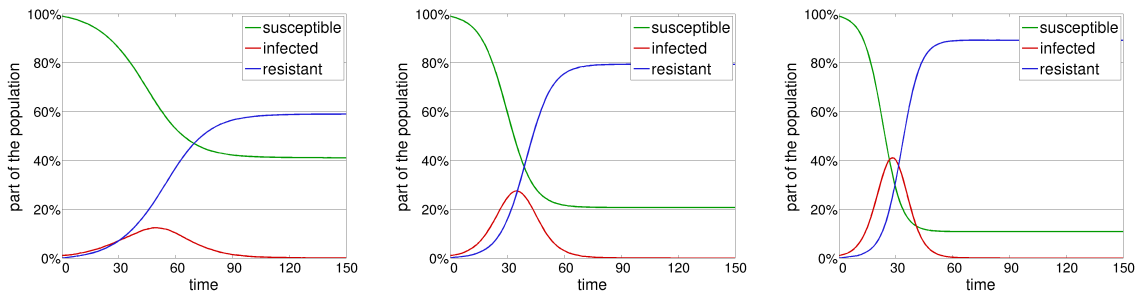
An SIR disease is commonly characterized by the time and the height of the peak, the duration until extinction and the remaining susceptibles after extinction. The progress of the disease depends on the initial values and on the parameter settings. Like an SIS epidemic, different modeling approaches might lead to different results caused by non-equivalent structures.

Figure 3.4 shows three typical results of SIR models.

### 3.3.3 Extensions of SIS and SIR models

Simple SIS and SIR models simulate a homogenous population where all persons are equally parameterized. Increased complexity might be realized in various ways:

- Consideration of persons or groups of persons with different disease parameters.



**Figure 3.4:** Typical results of an SIR model

- **Example:** Definition of persons with a weak immune system who have a higher transmission probability and a longer recovery time.
- **Example:** Add the age to the attributes of persons and define age-dependent disease parameters.
- Consideration of a more complex disease progression.
  - **Example:** Take an SIR epidemic where resistant persons become susceptible after a while so that they can be infected again. Such a disease is called SIRS.
  - **Example:** Distinction of infected persons (=I) in a latent state (=E) and an infectious state (=I). Such models are called SEIS or SEIR.
- Further complexity can be gained by any combination of the two methods described above.

## 3.4 Top-Down & Bottom-Up

Top-down approaches are models that simulate aggregated values, so they see the reality virtually from above.

In contrast, bottom-up approaches are models that simulate single individuals or particles with certain rules. Due to the rules of the individuals or particles, a behavior of the overall system emerges which is usually of interest.

The following chapters describe different modeling approaches that present schematically possible ways to model epidemics. Advantages and disadvantages of these approaches are also subject of discussion.

## 3.5 Markov models

Markov models are typical representatives of top-down approaches. They are popular in many areas due to their relatively simple handling [5].

### 3.5.1 Theoretical basics

**Stochastic processes:** To put it simply, a stochastic process is a family of random variables  $\{X_t, t \in T\}$  that are indexed by a totally ordered index set  $T$ .  $T$  usually describes the time but, for example, it can also describe the place.  $X_t$  represents the the system in time  $t$ . The set  $S$  of all possible values of the random variables is named value space or state space. The realization of a random variable represents its the observation and yields an element of the state space. A realization of the stochastic process, which consists of the realizations of all  $X_t$ , is called “path”. This means, a path is a family of states indexed by  $T$ . If  $T$  is a finite or a countably infinite set, then the process is called time-discrete and can be denoted as  $\{X_0, X_1, X_2, \dots\}$ . If  $T$  is an interval, then the stochastic process is called time-continuous. [47, p. 41-42]

**Markov processes and Markov chains:** Markov processes are special stochastic processes where the state in a time can only depend on most recent state but not on other states. This can always be tested since  $T$  is a totally ordered set. For Markov processes  $T$  is usually referred to as “time” and not any more “index set”. If the state space is discrete, then a Markov process is called “Markov chain”.

A Markov process is defined as follows: For any set of  $n$  time points  $t_1 < t_2 < \dots < t_{n-1} < t_n$ , the random variable  $X_{t_n}$ , for given values of  $X_{t_1}, \dots, X_{t_{n-1}}$ , depends only on  $X_{t_{n-1}}$ , the most recent known value. This property is called “memoryless” or “Markov property”. [43, p. 188-189]

The Markov property can be expressed explicitly for time-discrete Markov processes. Let  $P$  denote the probability and  $T = 0, 1, 2, \dots$  :

$$P[X_{t+1} = s_{j_{t+1}} | X_t = s_{j_t}, X_{t-1} = s_{j_{t-1}}, \dots, X_0 = s_{j_0}] = P[X_{t+1} = s_{j_{t+1}} | X_t = s_{j_t}]$$

This means that the probability that a system in state  $s_i$  in time  $t$  will be in state  $s_j$  in time  $(t + 1)$ , depends only on last the state and the time  $t$ .

**Remark:** Sometimes it is considered that states of a Markov model can depend on the states of the last  $m$  time steps. For discrete time, this property can be stated explicitly:

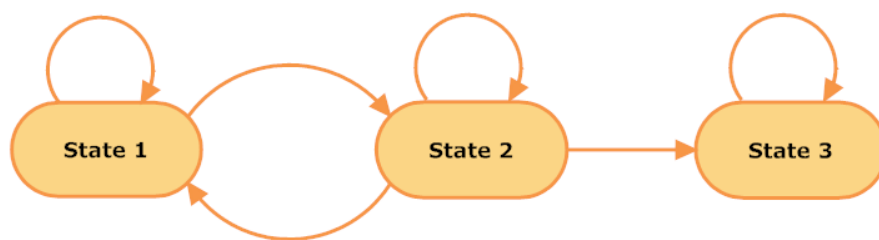
$$P[X_{t+1} = s_{j_{t+1}} | X_t = s_{j_t}, X_{t-1} = s_{j_{t-1}}, \dots, X_0 = s_{j_0}] = P[X_{t+1} = s_{j_{t+1}} | X_t = s_{j_t}, \dots, X_{t-m+1} = s_{j_{t-m+1}}]$$

If time and state space are discrete, then such stochastic processes can be named “Markov chains of order  $m$ ”. This approach is only partly documented in literature and rarely applied. Hence, it is not further used in this work.

**Visualization of Markov chains:** State transition diagrams are suitable to visualize Markov chains. In this diagram, all states are positioned somewhere on the figure, and

arrows between states show where transitions between states are possible. Self-referring arrows are possible but not necessarily needed. Without further statements, it is usually presumed that the system can remain in a state. Description of arrows can contain the probability that a system changes from one state to another in the following time unit. Hence, the sum of the probabilities of all outgoing arrows of a state is not more than 1 and, if it is less than 1, then system remains in this state with the remaining probability. Figure 3.5 shows such a state transition diagram.

Markov cycle trees, which are mainly used for Markov models, are another visualization technique which will be explained later.



**Figure 3.5:** Example of a state transition diagram

**Realization of Markov chains:** Markov chains can be realized in a descriptive way by matrices and vectors if they meet three requirements:

- The state space  $S$  is finite.
- The time  $T$  is discrete.
- The Markov chain is memoryless.

Then, the state space can be explicitly written as  $S = s_1, s_2, \dots, s_m$  with  $m \in \mathbb{N}$ . The transition matrix  $M = (p_{i,j})$  is an  $m \times m$ -matrix that holds:

$$P[X_{t+1} = s_j | X_t = s_i] = p_{i,j}$$

This means that  $p_{i,j}$  is the probability that a system that is in state  $s_i$  will be in state  $s_j$  in the following time point. From this it follows:

$$0 \leq p_{i,j} \leq 1 \text{ for all } 1 \leq i, j \leq m$$

The row sums of the matrix must equal 1:

$$\sum_{j=1}^N (p_{i,j}) = 1 \text{ for all } 1 \leq i \leq N$$

A transition matrix can either be constant for the whole time or be dependent on the time point  $t$ . [47, p. 163-164] [52, p. 25, p. 81-184] [45] [5]

**Distribution of a random variable:** Further calculations require introduction of a new term: The “distribution of a random variable” is a distribution on the state space which indicates the probabilities that a certain state is reached. For a discrete state space, the distribution of a random variable in  $s$  gives the probability that a realization of the random variable results in the state  $s$ . For a finite state space, this distribution can be interpreted as a vector, and its elements sum up to 1.

This definition can now be used for Markov chains. Let  $\underline{a}(t) = (a_1(t), a_2(t), \dots, a_n(t))$  be the distribution of the random variable in time  $t$ . Then,  $a_i$  indicates the probability that the system is in state  $s_i$  in this time point.  $\underline{a}(0)$  is the distribution of the first time point, hence it is called initial distribution or initial vector. Furthermore, one has  $\underline{a}(t+1) = \underline{a}(t) \cdot M$  for constant transition matrices respectively  $\underline{a}(t+1) = \underline{a}(t) \cdot M(t)$  for time-dependent transition matrices. This statement is intuitively true and easy to prove.

Using this shown concepts, it is possible to write down time-discrete Markov chains with finite state spaces, calculate realizations and immediately specify the distribution in every time step.

### 3.5.2 From the theory to the model

**Theory:** The previous chapter defined stochastic and Markov processes as given families of random variables. Based on this knowledge, it is possible to define Markov models. We consider a dynamic system which is evaluated in equidistant time points  $0, 1, 2, \dots$ . In every time point, the system is in one of the states  $s$  of a finite amount of states  $S$ . The probabilities  $p_{i,j}(t)$  indicate the probability that a system in state  $i$  in time  $t$  will be in state  $j$  in time  $(t+1)$ . These probabilities can be considered as a transition matrix that is either constant or time dependent. Due to given transition probabilities, the progression of the states fulfills the Markov property. Hence, the constructed system is a Markov chain.

**Remark:** Markov models can be extended to “Markov decision models” by introduction of a finite set of actions or decisions  $d \in D$ . In every time point, an action is performed that affects the transition probabilities. Hence, they are denoted as  $p_{i,j}(a)$  for the time-invariant case. Actions cause costs which are, additionally to the state of the system in time  $t$ , of interest.  $c_i(d)$  indicates the costs of an action  $d$  whenever a system is in state  $i$ . This makes it possible to define a time-dependent cost function which sums up the costs of single actions in every time point. This approach is commonly used, but it is not needed in this work because this study does not consider economic evaluations. [52, p. 234-237] [51]

**Realization of Markov models for health economics:** The following deliberations are generally based on the results of the ISPOR-SMDM Joint Modeling Good Research Practices Task Force [51]. Markov models in the field of health economics are used to



simulate the temporal progress of persons or of a population. Two model types as applications of Markov chains are commonly used: Microsimulation and cohort models. Both approaches are based on the same Markov model but they perform different calculations. The underlying Markov model is defined for single persons as follows:

- The time is given as a sequence  $0, 1, 2, \dots, n$ . The number of simulated time units and the corresponding real time of one time unit should be chosen case-related.
- The states should cover all relevant properties of a person and should be mutually exclusive.
- The transition probabilities must be given for all possible transitions between states. They can be constant for the whole simulation time or time-dependent. Convenient notations are transition matrices or Markov trees which will be explained later.

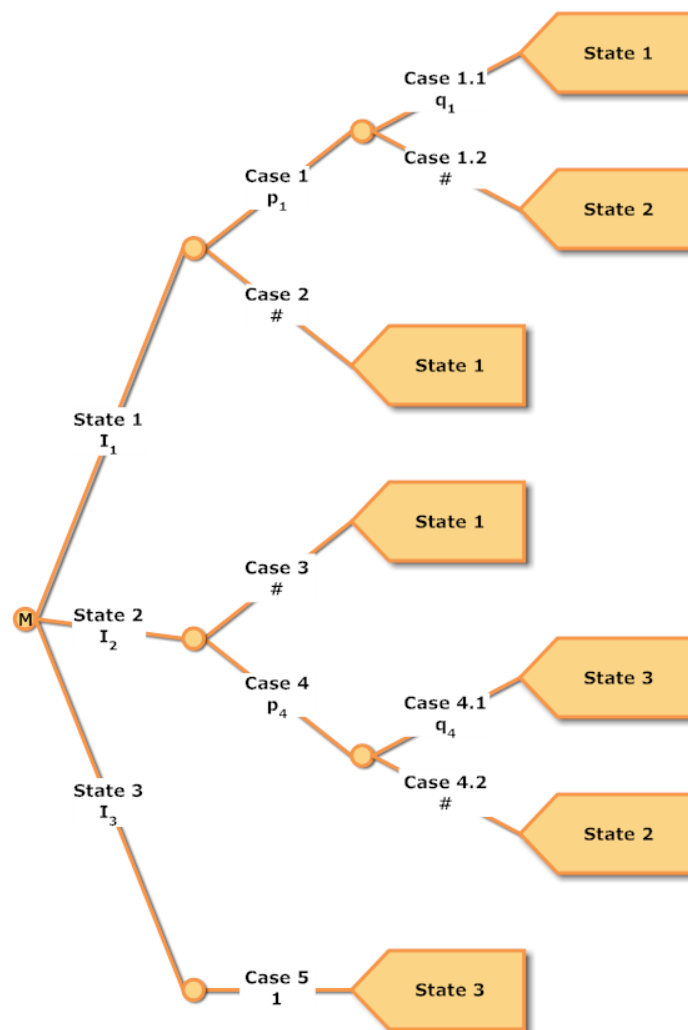
This structure builds a model which is able to simulate the state progression of a person over time.

Often, persons are not equally parameterized. Instead, they differ by properties that have an influence on the questioned states. For example, the progression of a disease might depend on the age of the person. Furthermore, Markov models are commonly used to analyze interventions by comparison of different experimental and control groups. One is advised to split up the population by properties and interventions at first and then define a Markov model for each group which only models the questioned circumstances.

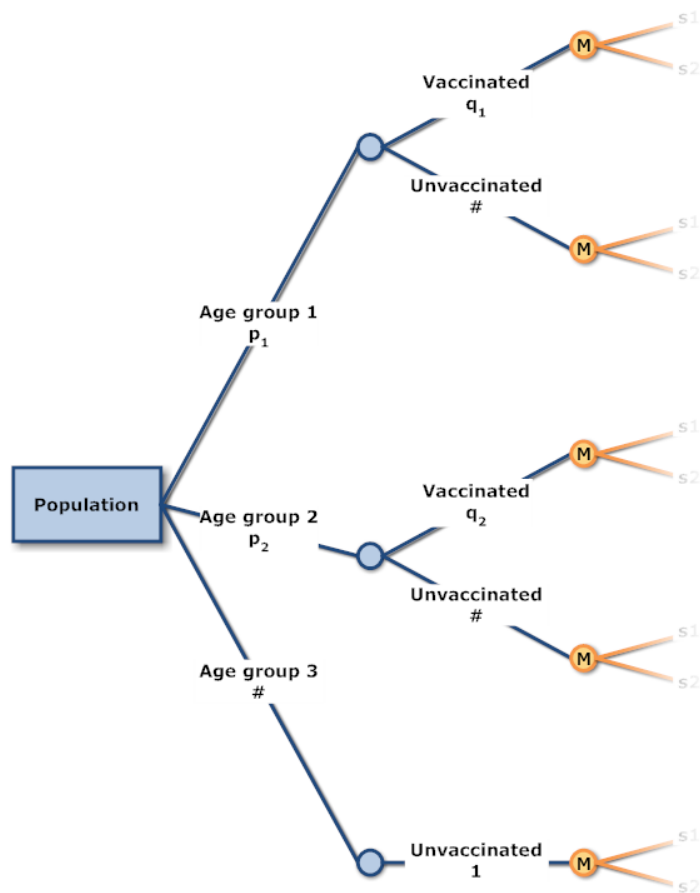
**Visualization of Markov models:** State transition diagrams are a visualization technique which has been explained already. Trees are an alternative technique. These trees are decision trees, which means that every branch has a probability. The probabilities of all outgoing branches of a single node must sum up to 1, hence the probability of one of the outgoing branches can be written as  $\#$  to represent the remaining probability.

These trees consist of two classes of decision trees. The first part is a tree which describes the splitting of the population into the mentioned groups. The leaves of this tree are Markov nodes which are the root nodes of another decision tree, called Markov tree, which visualizes the Markov model of this group. The Markov node has one outgoing branch for each state of the state space. The description of these edges usually contains the initial distribution. These nodes are followed by subtrees that represent cases and their probabilities. Every leaf of these subtrees contains the state that has been reached along this specific path of branches. With this technique, the transition probabilities can be constructed in a descriptive way and do not need to be given as single values. Figure 3.6 shows a Markov tree and figure 3.7 shows a decision tree for splitting the population with the attached Markov trees.

Based on the preliminary work it is now possible to describe the two simulation methods.



**Figure 3.6:** Example of a Markov tree. The variables  $I_j$  represent the distribution of the initial state and  $p_j$  and  $q_j$  show the probabilities for cases.



**Figure 3.7:** Example of a decision tree. The variables  $p_j$  and  $q_j$  show the probabilities for decisions.

**Microsimulation:** Microsimulation, also called Monte Carlo simulation, is a technique that simulates single persons whose states are simulated by a Markov model. The significance of simulation of single persons is usually low, hence it is required to simulate enough persons of each group to make useful statements about expectation values and variation.

**Cohort simulation:** A cohort is a group of the population. A cohort simulation calculates the expected distribution of each group for each time point that arises when every person runs its own Markov model. These values correspond to the distribution of a Markov chain, hence they can be calculated by vector-matrix-multiplications. The distribution of the whole population can be reached by multiplying the group distributions with the group size and summing them up. Cohort simulations require less effort than Microsimulation and often, they already yield the desired results of the Markov model.

**Used terms:** Such models are often referred to as “state transition models”. This term is not exactly specified, it just describes a class of models where something switches between available states over time. Figure 3.8 illustrates relationships between commonly used terms in this area of modeling.

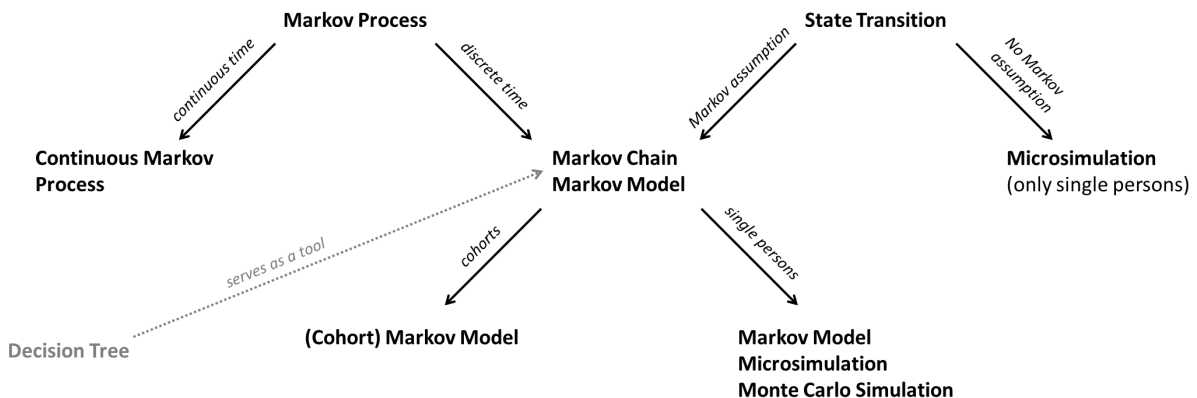


Figure 3.8: Commonly used terms for models in the area of stochastic processes

### 3.5.3 Discussion

**Avoiding the Markov property:** The Markov property means that a system can only depend on the most recent state, no matter which path it took to reach this state. Technically, it is possible to avoid this Markov property if a state contains other past states.

One approach is to distinguish between system states and actual states. Then the system state in time  $t$  is a vector that contains the actual state in time  $t$  and the past  $m$  actual states. For calculation of the system state in time  $(t + 1)$ , the last element of the vector, which contains the furthestmost state, is dropped and the new actual state is calculated

and added. In case of finite state spaces, this means that a Markov chain of order  $m$  can always be described as a memoryless system.

Another possibility is to define a state in a way that it contains the past state. This means that the realization of  $X_1$  contains the realization of  $X_0$ . Then the realization of  $X_2$  contains the realization of  $X_1$  which still contains the realization of  $X_0$ . Hence a state contains the whole history of states in a recursive way.

Both cases do not affect the Markov property. However, the state space and the number of transition probabilities grow exponentially because a state can contain all possible combinations of past states. Hence, it can be very complicated to build such a system. An implementation by appropriate vectors and functions seems possible, but constructing a Markov models with transition matrices or Markov trees is almost impossible due to the fast growing number of states.

**Feedback loops:** Many systems and models are able to produce feedback loops. This means that single components interact with each other and produce a system behavior that would not be possible without this interaction. Markov models, in contrast, calculate the transition probability for each person due to their construction. Therefore, they do not allow feedback loops. But, it is an important question if there can be stochastic processes with feedback loops. It is clear that mutual influence generally does not need information from the past so it does not violate the Markov property. But, it is subject for research if such a system can be a valid stochastic process. It is easy to build a process with feedback loops. Following the idea of the Markov model, one can recalculate the transition matrix in every time point, depending on the states of other persons. Mathematically speaking, this means that the transition Matrix of a Markov chain in time  $t$  depends on the distribution of the system in  $t$ . The realization of the system is still subject to variations due to randomness and it can be evaluated in each time point. Thus, it is an ordered family of random variables and hence it is a stochastic process.

**Remark:** Definitions of a stochastic process often require that all random variables  $\{X_t, t \in T\}$  are defined on the same probability space [55, p. 1]. This definition raises several questions. If the probability space represents a single event (for example a coin toss), then the realization of a random variable can only be the result of this single event. Hence, it is impossible to include information of past states in the random process. This means that neither typical Markov processes nor the common example “random walk” are stochastic processes any more.

Alternatively, it is possible to build the probability space in a way that it contains the whole stochastic process. This means that the whole path of the process can be calculated by a single element of the sample space. In other words, an element of the sample space allows calculation of the realization of  $X_t$  for every  $t \in T$ . In this case the probability space can be constructed arbitrarily so that any desired feedback loop in a Markov process can be represented.

### 3.5.4 SIS epidemics with Markov models

An SIS model needs two states:  $s_1 =$  susceptible and  $s_2 =$  infected. These states are sufficient to characterize the health state of a person in an SIS model.

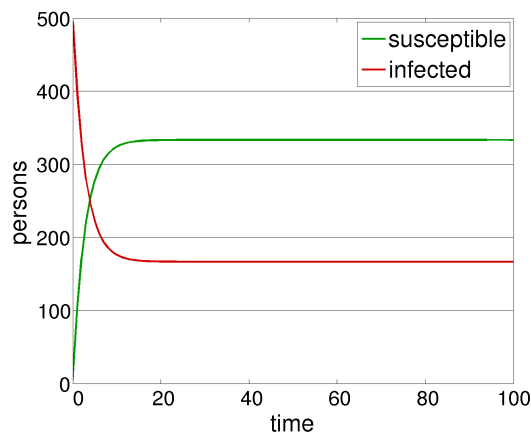
An example of a transition matrix is:

$$M = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$$

This means that the probability for a susceptible person to get infected in a time step is 10% and for an infected person to get susceptible is 20%. This is equivalent to the statement that 10% of the susceptible persons in the population get infected and that 20% of the infected persons get susceptible per time step.

Let  $(0.99 \ 0.01)$  be an exemplary initial distribution of the system. Then 1% of the population is infected while the rest is susceptible in the beginning.

A cohort simulation of the model with the given parameters shows a system that reaches an equilibrium after a few time steps (figure 3.9).



**Figure 3.9:** Result of a typical SIS Markov model

### 3.5.5 SIR models with Markov models

An SIR model needs three states:  $s_1 =$  susceptible,  $s_2 =$  infected and  $s_3 =$  resistant.

A transition matrix of such a model looks like this:

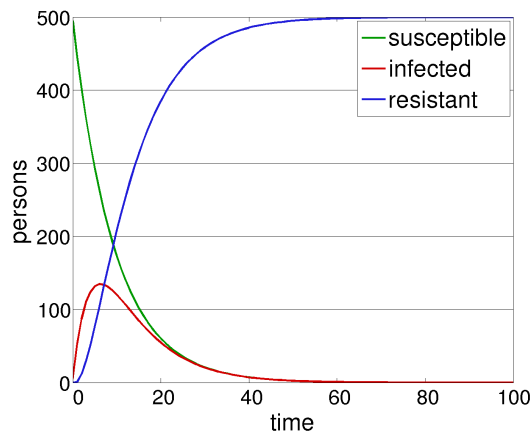
$$M = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0 & 0.8 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$$

This means: The probability that a susceptible person gets infected in a time step is 10% and the probability that an infected person recovers and becomes resistant is 20%. Again this is equivalent to the statement that in a population 10% of the susceptible persons

get infected and 20% of the infected persons recover per time step.

Let the initial distribution of the system be  $(0.99 \ 0.01 \ 0)$ . This means that 99% of the initial population is infected and the rest is susceptible.

The cohort simulation with the given parameters shows an epidemic where the number of infected people in the population increase rapidly in the beginning, reaches a maximum and then slowly decreases to zero (figure 3.10).



**Figure 3.10:** Result of a typical SIR Markov model

**Remark:** It is important to say that Markov models are hardly used for modeling propagation of epidemics. Instead, they are used to simulate sickness numbers over long time periods without consideration of the concrete dynamics of the epidemic. The following chapters in this work explain in detail why the Markov approach is not able to represent the spread of epidemics correctly.

## 3.6 Differential equations

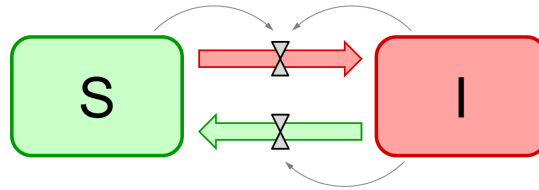
Differential equations are one of the oldest and best examined modeling approaches for epidemics. Since they calculate with cumulative values and not with single persons, they are top-down approaches.

The general, extensive theory of differential equations is not explained here. Instead, let me refer to a huge selection of books.

### 3.6.1 SIS models with differential equations

An SIS model leads to a system of two equations (formula 3.1) which can also be visualized (figure 3.11).

$$\begin{aligned}
 \dot{S}(t) &= -\alpha \cdot I(t) \cdot S(t) + \beta \cdot I(t) \\
 \dot{I}(t) &= \alpha \cdot I(t) \cdot S(t) - \beta \cdot I(t) \\
 &\text{with the initial values} \\
 I(0) &= I_0 \geq 0, \quad S(0) = N - I(0)
 \end{aligned}
 \tag{3.1}$$



**Figure 3.11:** Visualization of the differential equation SIS model

With:  $S(t)$  ... Number of susceptible persons in time  $t$   
 $I(t)$  ... Number of infected persons in time  $t$   
 $\alpha, \beta > 0$  ... Model parameters  
 $N$  ... Total number of persons

This differential equation has an exact solution. Since  $N$  is constant and no persons get lost,  $S(t) = N - I(t)$  is always true. As a consequence, it is enough to solve  $I(t)$  explicitly and then calculate  $S(t)$  with it. We plug in for  $S(t)$  in  $\dot{I}(t)$ :

$$\dot{I}(t) = \alpha \cdot \left( N - \frac{\beta}{\alpha} - I(t) \right) \cdot I(t)
 \tag{3.2}$$

For simplification we define  $D := N - \frac{\beta}{\alpha}$ . Then two cases have to be solved separately:

**Case 1:**  $D = 0$

$$\dot{I}(t) = -\alpha \cdot I^2(t) \Rightarrow I(t) = \frac{I_0}{\alpha \cdot t \cdot I_0 + 1}
 \tag{3.3}$$

**Case 2:**  $D \neq 0$

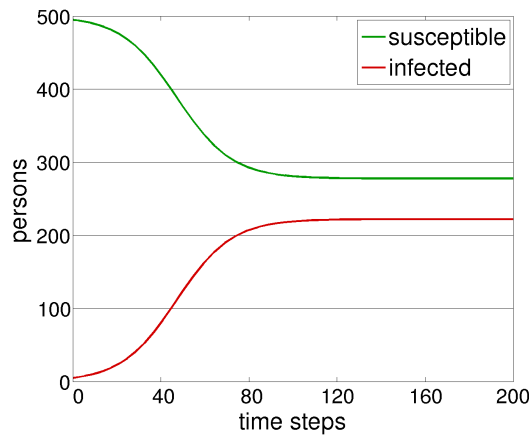
$$\dot{I}(t) = \alpha \cdot (D - I(t)) \cdot I(t)
 \tag{3.4}$$

Rearrangement of the equation, indefinite integration and plug in of the initial values yields:

$$I(t) = \frac{I_0 \cdot e^{\alpha D t}}{1 - \frac{I_0}{D} (1 - e^{\alpha D t})}
 \tag{3.5}$$

Such an epidemic model always leads to an equilibrium. In case of  $D \leq 0$  the epidemic becomes extinct - then the equilibrium means that all persons are susceptible. The case  $D > 0$  leads to a real equilibrium and  $I$  equals asymptotically  $D = N - \frac{\beta}{\alpha}$  (by application of L'Hôpital's rule). Figure 3.12 shows a typical result of such a differential equation SIS model.





**Figure 3.12:** Typical result of a differential equation SIS model

### 3.6.2 SIR models with differential equations

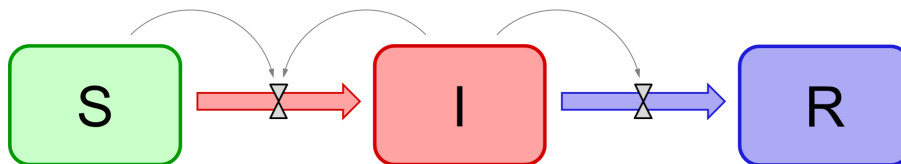
A SIR model consists of three population groups and therefore requires three equations as shown in formula 3.6. Figure 3.13 shows an illustration of this approach.

$$\begin{aligned}
 \dot{S} &= -\alpha \cdot I(t) \cdot S(t) \\
 \dot{I} &= \alpha \cdot I(t) \cdot S(t) - \beta I(t) \\
 \dot{R} &= \beta \cdot I(t)
 \end{aligned}
 \tag{3.6}$$

with the initial values

$$S(0) = S_0 \geq 0, I(0) = I_0 \geq 0, R(0) = N - S(t) - I(0)$$

With:  $S(t)$  ... Number of susceptible people in time t  
 $I(t)$  ... Number of infected people in time t  
 $R(t)$  ... Number of resistant people in time t  
 $\alpha, \beta > 0$  ... Model parameters  
 $N$  ... Total number of persons



**Figure 3.13:** Visualization of the differential equation SIR model

This system of differential equations is not exactly solvable, in contrast to the SIS model. At least the theory states that this system has a solution and that it is unique for given initial values  $S(0)$ ,  $I(0)$  and  $R(0)$  (referring to the Picard–Lindelöf theorem [3, p. 68-71]). Figure 3.14 shows a typical result of an SIS model of differential equations.

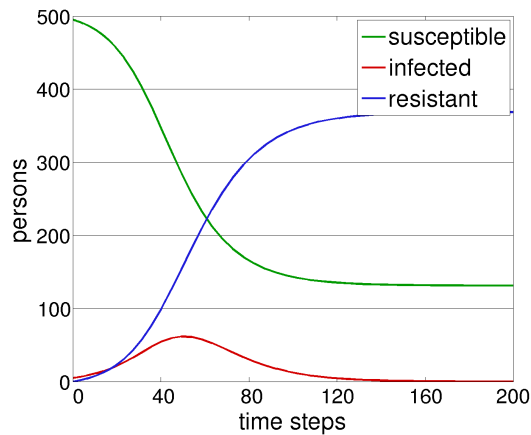


Figure 3.14: A typical result of an SIS model of differential equations

### 3.7 Cellular Automata

The concept of cellular automata (=CA) was introduced in the 1950s by John von Neumann [56]. Based on this idea, John Conway developed the well known “Game of Life” [18]. Stephen Wolfram showed in his books “Theory and Applications of Cellular Automata” and later in “A new kind of science”, how complex behavior can be achieved upon simple rules in cellular automata [60] [61]. Based on the basic ideas, several types of cellular automata have been developed and classified for different fields of applications [17].

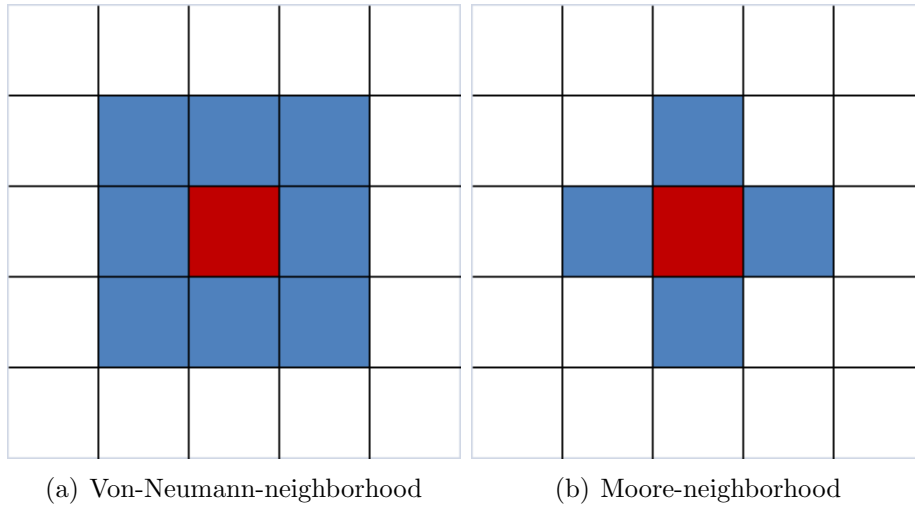
Cellular automata are constructions for calculation of spatial dynamic behavior of systems. They consist of a finite number of cells and states, while a specific cell is always in exactly one of those states in a point in time. The automaton is calculated time-discrete, that means time point by time point. The state of a cell in time  $t + 1$  only depends on its own state and the states of its neighboring cells in time  $t$ .

Hence, a cellular automaton is mathematically defined by these elements [10]:

- A space  $R$  (cell space)
- A finite neighborhood  $N$
- A set of states  $Q$
- A local transition function  $\delta : Q^N \rightarrow Q$

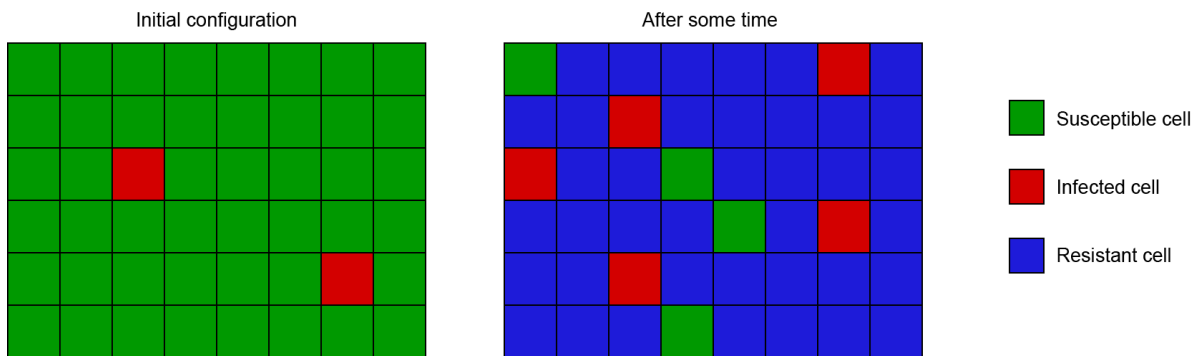
In the most simple case a cellular automaton consists of a 2-dimensional grid and square-shaped cells. For that case two important neighborhoods exist: The Moore- and the von-Neumann-neighborhood (see figure 3.15) [59, p. 28].

For epidemic simulations, the different infection states build the set of states. For example, susceptible cells get infected if at least  $x$  cells in their neighborhood are infected.



**Figure 3.15:** Neighborhoods on cellular automata

Afterwards, they pass through a number of infection states until they get susceptible (in an SIS model) or resistant (in an SIR model) again (see figure 3.16). Fuentes and Kuperman compared such a system in 1999 with a differential-equation-model [15].



**Figure 3.16:** Realization of an SIR model as a cellular automaton

An extension of cellular automata are particles that move around on cells of a cellular automaton or on a grid. These approaches are described and examined very well under the terminology “Lattice Gas Cellular Automata” (LGCA), which are mainly used for simulation of fluids and gases [59, p. 51-57]. For epidemic simulations, this approach is extended again so that particles have an infection state and can infect other particles in their neighborhood. This methodology has been introduced by Yakowitz 1990 [63]. An epidemic model on a hexagonal LGCA is schematically shown in figure 3.17. Fukš and Lawniczak performed extensive technical analyses on models based on this idea [16]. Further model comparisons with differential equations have been shown by Štefan Emrich and Günter Schneckenreither [65] [50].

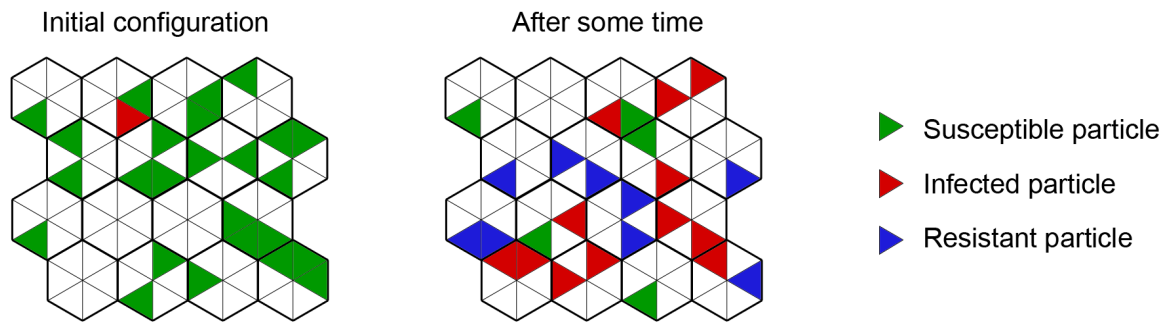


Figure 3.17: Realization of an SIR model based on an LGCA

## 3.8 Agent based models

Agent based models have already been defined and described in chapter 2. Based on the general ideas, this chapter describes agent based models for epidemic simulation.

Such a model generally consists of three parts:

1. **Agents:** They represent persons.
2. **Contacts:** Contacts happen between persons.
3. **Disease:**
  - **Infection:** A susceptible agent can get infected upon a contact with an infected agent.
  - **Recovery:** An infected agent recovers after some time. This does not depend on contacts with other agents.

### 3.8.1 Example of an agent based epidemic model

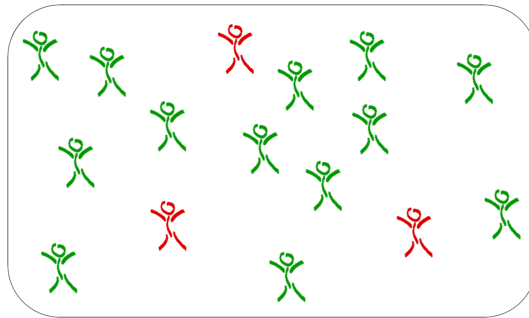
- **Agents:** They represent persons.
- **Attributes:**
  - **Infection state:** An SIS model allows two values: S and I. And SIR model allows three values: S, I and R. The states are visualized in figure 3.18.
- **Rules:**
  - **Contact rule:** An agent has a certain number of contacts with other agents within a time interval (figure 3.19).  
**Parameter:** Number of contacts per time interval.

- **Infection rule:** If the agent is susceptible and has a contact with an infected agent, then it changes its state with a certain probability to I - an infection happens (figure 3.20).

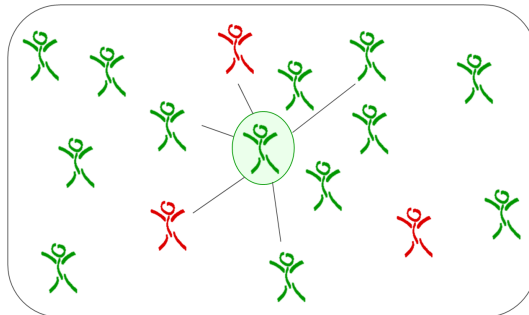
**Parameter:** Infection probability.

- **Recovery rule:** Infected agents change their infection state after a certain time interval to S (in an SIS model) or to R (in an SIR model) - this is called recovery (figure 3.21).

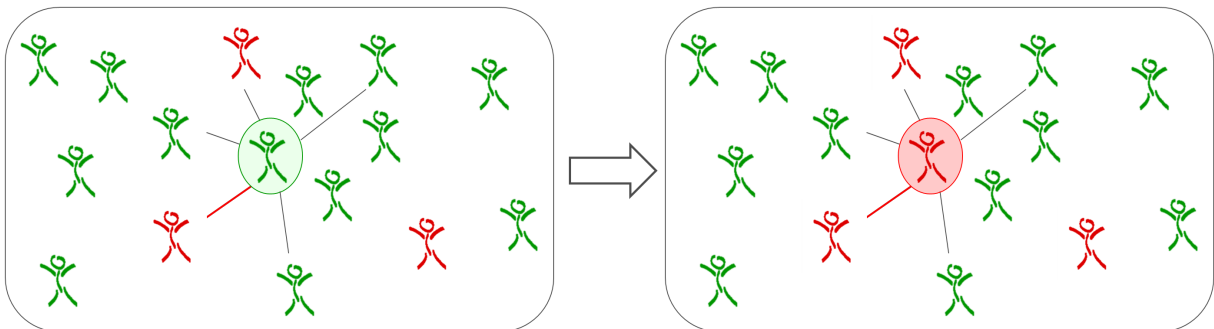
**Parameter:** Time to recovery (=recovery time).



**Figure 3.18:** Agents in an epidemic model

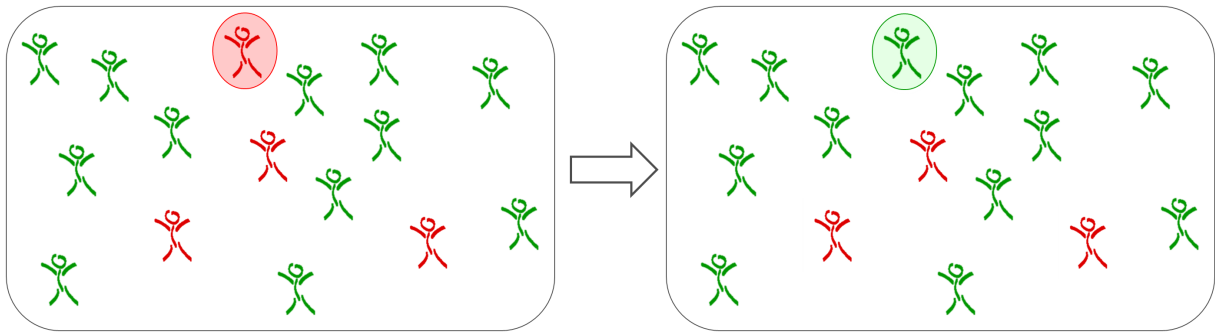


**Figure 3.19:** Contacts between agents in an epidemic model



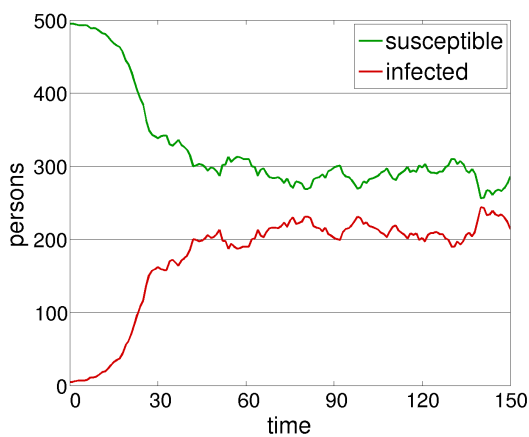
**Figure 3.20:** Infections of agents in an epidemic model based on contacts

**Results:** The figures 3.22 and 3.23 show results of a simulation with an agent based epidemic model with 500 agents of an SIS epidemic and an SIR epidemic. A random noise on the graphs is clearly visible. It is based on the fact that agents do not behave deterministically. Instead, every agent underlies random decisions based on probabilities

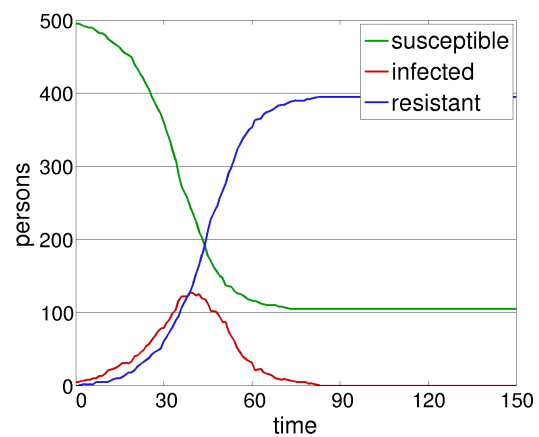


**Figure 3.21:** Recovery of agents in an epidemic model

all the time. Hence, in some time steps, more or less agents than the statistic mean gets infected. Comparison of these curves with the differential equation model clearly shows this effect as well (compare to figures 3.12 and 3.14).



**Figure 3.22:** Result of a typical agent based SIS model



**Figure 3.23:** Result of a typical agent based SIR model

### 3.8.2 Extension of agent based epidemic models

An important benefit of agent based models is their flexibility and extensibility. Possible extensions are presented below, based on the model in chapter 3.8.1.

**Parameter changes:** Every agent is parameterized individually. In the most simple case all agents have same parameters. But in other cases, some agents can get assigned higher contact rates, lower infection probability or a longer recovery time without any structural changes on the model and even without affecting other agents.

**Additional attributes:** Agents can get new attributes. Their implementation is almost without effort because they do not affect any other parts of the model. Useful application

of new attributes is to consider them in some agent's rules which usually is simple to implement, as stated in the next paragraph.

Persons in an epidemic model could be given additional attributes like age, gender or strength of the immune system.

**Adding/change of rules:** Rules of agents can be changed or added individually. No changes in the rest of the model have to be done as long as these rules do not require circumstances that do not exist in the model yet. For example, this means that an agent can only fly to the moon if the moon exists in the model. This also allows to redefine, change, split up or combine classes of persons without much effort.

Making use of this idea for epidemic models, contact rates could depend on the age or the infection probability could depend on the immune system of the person.

**Conclusions:** The example shows the benefit of agent based modeling very well, which is already generally stated in chapter 2.5. Because of the simplicity of these kind of modeling, one should not forget that changes of parameters or rules of a few agents might already have strong impact on the system behavior and the result. Hence, an extensive validation is very important for justification of working with a valid, correct model.

## 3.9 Comparison of the presented approaches

Chapter 3 deals with four important modeling approaches which are commonly used for simulation of the spread of epidemics. In this context, the question arises which the "accurate" or at least the "most accurate" model is.

First of all, it needs to be cleared that a model is always a simplification of the reality. The actual problem, the concrete question and also the knowledge (and available data) on the real system should be considered in choosing the right approach.

All presented modeling approaches have its justification for simulation of epidemics. It depends on the actual situation which one should be used for modeling and simulation of a concrete problem. Therefore, advantages and disadvantages of these models are presented below.

### 3.9.1 Markov models

#### Advantages

- Direct extrapolation of data and hence relatively simple parameterization.
- Transparent calculation that is clear even for non-modelers.

- Programming is not necessary. Markov models can even be calculated with Microsoft EXCEL. However, there is specialized software which still require programming knowledge.

### Disadvantages

- There are no feedback loops because the models work with static transition probabilities (for example, susceptible persons can be infected even though no infectious persons exist). Hence, possibly existing dynamic effects are ignored.
- Temporal dynamics are generated by given, time-dependent transition matrices but they are not based on structural dependencies.
- There are limited level of details. Every consideration of a detail requires splitting of at least one new cohort or adding at least one new state. Hence, more and larger transition matrices appear and the system is much harder to parameterize.

### 3.9.2 Differential equations

#### Advantages

- Extensive mathematical theory for solution and analyses of differential equation models.
- Possibility of descriptive visualization by system dynamics.
- No explicit consideration of details and relationships on a micro-level, hence they do not need to be known.

#### Disadvantages

- The model itself is not descriptive and hard to understand for non-specialists.
- Common usage of abstract parameters that can only be set by calibration and that are often not directly related to the real system.
- Very limited level of details. Every additional consideration of a detail requires at least one additional variable and an equation. Furthermore, the other equations are usually changing. This leads to very large systems of equation which are hard to calibrate and to solve.

### 3.9.3 Cellular automata

Cellular automata can be seen as a prestage of agent based models where single components with strict and fixed rules are used.



### Advantages

- Relatively simple description of the system.
- Very standardized way of modeling.
- Dynamic “emergent behavior” happens obviously and clearly, which is described in chapter 2.5.
- Simple visualization.

### Disadvantages

- Same type and parameterization of all cells.
- Cellular automata are often a high abstraction level of the real system.
- No mathematic theories for analysis of behavior and results.

### 3.9.4 Agent based models

#### Advantages

- **Clarity also for non-specialists:** The epidemic in the model is based on the same reasons as in the real world: transmissions from person to person.
- **Direct usage of known data as parameters and rules:** The contact behavior can be observed in social studies. The disease states and progressions are generally medically known.
- **Flexibility of the system:** Contact behavior and disease progressions can be changed easily without need to adopt any other parts of the model. This allows efficient testing of fictional or real assumptions, but it also increases reusability of the models in other projects.
- **Easiness of creation of complex system dynamics that do not need to be known priorly (“emerging behavior”):** Different parameterizations and interventions (for example vaccinations) usually result in a nonlinear change of the disease spread.  
**Remark:** The following chapters examine such effects which can be described very well by appropriate definitions of herd immunity and serotype shift.
- **Flexible scalability:** The simulated number of persons but also personal attributes and rules are generally extensible as desired. The only restrictions are memory and runtime on the computer.

### Disadvantages

- Simulation of details might require extensive validation because their impact on the whole system is often a priori unknown. Especially the occurrence of an unexpected epidemic spread by certain assumptions must be examined carefully and meticulously to find out whether the results are based on wrong or insufficient assumptions or if they are indeed modeled correctly.
- A large number of agents require powerful computers and long runtimes. Efficient coding and appropriate scaling based on prior knowledge helps to overcome this problem.
- There are no standardized methods for analysis of the system behavior and of the results. The idea of agent based modeling does not provide any methods on how the spread of epidemics might be recorded and described.

**Remark:** The following chapters show that different situations often require different statistical analyses on persons or on parts of the population.

# Chapter 4

## Herd immunity: Terminology, motivation and origin

### 4.1 Usage of the term “herd immunity” in the literature

Herd immunity is an effect which is of importance for interventions against epidemics.

The usage of the term “herd immunity” in the literature is very divergent; exact explanations and specifications are rarely found.

The term “herd immunity” occurs for the first time in 1923 in a paper titled “**The spread of bacterial infection: the problem of herd immunity**” by the British bacteriologists William Whiteman Carlton Topley and Graham Wilson. Based on experiments with mice, they describe in their study that it is not enough to take account of the immunity of single individuals. Additionally, the resistance of whole populations should also be considered. [53]

Paul E. M. Fine examines in 1993 in “**Herd Immunity: History, Theory, Practice**” the historic approaches to herd immunity and combines the ideas. He interprets herd immunity as a resistance of a population which describes if an epidemic breaks out or dies out, depending on given initial configurations and values. He also compares his theoretic conceptions to practical examples like the impact of vaccinations against smallpox or measles. [14]

Michael Haber et al. define a quantity called “herd immunity effectiveness” in their paper “**Herd immunity and pneumococcal conjugate vaccine: A quantitative model**”. This quantity indicates the relative reduction of sicknesses in a not vaccinated part of the population. [21]

T. Jacob John and Reuben Samuel distinguish between the terms “herd immunity” and “herd effect” in “**Herd immunity and herd effect: new insights and definitions**”.

Based on three other studies they use “herd immunity” for the part of immune people in a population [41, 42, 46]. In contrast to that, they denote with “herd effect” the change of epidemiologic parameters (for example the reduction of sicknesses) in the unvaccinated part of the population. [25]

“**Dorland’s Illustrated Medical Dictionary**” explains herd immunity as the resistance of a group against an attack of a disease, caused by immunity of a sufficient large part of the group, so that potential sick people do not meet enough susceptible people to induce an outbreak of the epidemic. [12]

Gerald T. Keusch and Kenneth J. Bart explain herd immunity in “**Immunization and Vaccine Use**” (**Harrison’s Practice of Internal Medicine**) as an indirect protection of not vaccinated or not immune persons because the pathogen cannot spread freely due to prevalence of immune persons. [31]

In “**Oxford Textbook of Medicine**” herd immunity is described as an effect which leads to extinction of a disease by high vaccination rates. [57, p. 353]

Herd immunity also plays a role in the **HTA report of the year 2008 to pneumococcal vaccinations in Germany**. At first, herd immunity is explained as resistance of a population so that a pathogen cannot spread and can no longer become an epidemic. Later on the term “herd immunity effect” is used to describe infection numbers in the unvaccinated population. [9, p. 56-58]

Similar effects are also explained in “**The impact of indirect (herd) protection on the cost-effectiveness of pneumococcal conjugate vaccine**” by using the terms “indirect protection”, “herd protection” and “community immunity”. [23]

Herd immunity is also divergently used in studies that examine the spread of epidemics:

Guy Katriel and Lewi Stone find out in their model that the number of infected people does not linearly depend on the infected people at simulation start. They justify this effect by herd immunity. [27]

Ana Perisic and Chris T. Bauch explain herd immunity as an effect, where unvaccinated people hardly get infected in case of high vaccination rates, even though they are still susceptible. [44]

## 4.2 Usage of the term “herd immunity” in this work

The following, general statement serves as a basis for the subsequent mathematical definition of herd immunity:

**Herd immunity is an effect in the area of infectious diseases that indicates a benefit for unvaccinated persons by vaccination of other persons.**

In this context a second definition of a term seems to be useful:

**Complete herd immunity indicates the state of a population (“herd”), when a particular disease cannot spread successfully and becomes extinct instead.**

### Remarks:

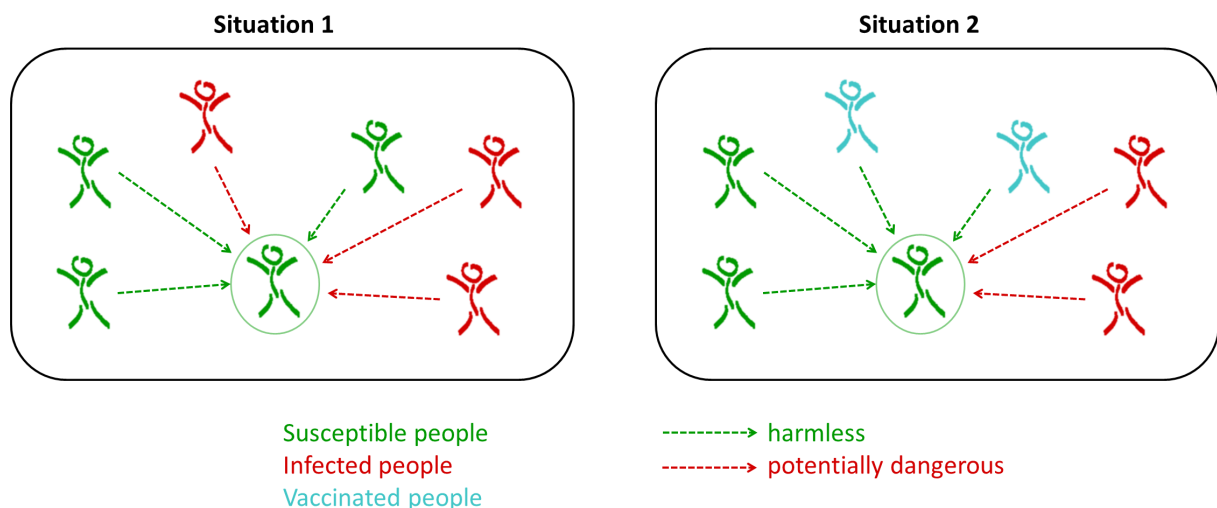
- The previous considerations show clearly that herd immunity is an effect which plays a role on diseases that are spread in a population by transmission. Other diseases that are not transmittable are obviously not affected. Therefore only models, that simulate infectious diseases, are important for examination.
- For consideration of vaccinations only persons that actually get immune by the vaccine need to be distinguished. Vaccinated persons that do not get immune play the same role as unvaccinated persons for the spread of a disease. From the health-care system’s point of view, such persons cause costs and need resources but these aspects are not relevant for examination of herd immunity.
- Depending on the disease, people can become resistant against the pathogen after recovery of an infection and cannot get infected again. For calculations of herd immunity it is important to distinguish between immunity due to an infection and immunity due to a vaccination because the focus of interest is the benefit of people who are not immunized by an intervention.
- Persons can also get resistant by taking drugs for a shorter or longer time instead of vaccinations. It is also possible to simulate such treatments with epidemic models and study the herd immunity.
- The benefit for unvaccinated people should be interpreted by values that describe the disease very well. These values could be, dependent on the point of view and on available information, newly infected people, carrier or sick people.
- Complete herd immunity is reached when enough people are vaccinated so that the disease cannot spread any more and becomes extinct. This is the highest possible benefit for unvaccinated people. This state is characterized by a maximal herd immunity. Higher values cannot be reached by herd immunity because a benefit beyond extinction of a disease is not possible.
- **About the term:** Objectively the usage of the term “herd immunity” in this work can be misleading because it does not indicate the immunity of a group. Instead, it describes the benefit of a not immune group due to immunity of other people. In this sense, I agree more with T. Jacob John and Reuben Samuel [25] that the term “herd effect” is terminologically more suitable. Yet I decided to call it “herd immunity” due to its common usage. Last but not least, “complete herd immunity” is not only closely related to “herd immunity”, it also represents its original meaning. So this work does not discuss something completely different from the original meaning.

### 4.3 Emergence of herd immunity

The basis is a transmittable infectious disease. Susceptible persons (persons who can get infected potentially) catch the disease with a certain probability from infectious people (=people who can spread the pathogen). The actual probability in reality depends of many factors and also varies from person to person. A crucial factor is the number of potentially infectious contacts of a susceptible person. These are contacts with infectious persons. Obviously, general personal health of the susceptible person, intensity and duration of the contacts and other influences also play a role for transmission.

**Situation 1:** This situation considers an infectious disease in a population where nobody is vaccinated. The probability of a person to get infected depends on the number of infectious people around this person. Figure 4.1 (left) visualizes this circumstance using red and green arrows.

**Situation 2:** Situation 2 refers to a state where a part of the population is vaccinated. All other aspects that are not illustrated (like social behavior or general health of the person) are assumed to correspond to situation 1. Then one can state that a susceptible person has less potentially infectious contacts compared to situation 1, and so they have a lower probability to get infected by someone. The schematic visualization shows this circumstance in a way that in situation 2 less red arrows go to the marked person compared to situation 1 (figure 4.1, right).



**Figure 4.1:** Visualization of the infection probability of the marked person.

**Situation 1:** No people are vaccinated.

**Situation 2:** A part of the population is vaccinated.

**Remark:** In reality, the actual infection probability in a case like situation 2 does not only depend on the number of vaccinated people, but also on a person's individual social behavior.

**Background:** The example of situation 1 and 2 shows only a very simplified view of this effect. Actually, it could happen that only susceptible people get vaccinated. Then, the remaining susceptibles have no benefit from this intervention at first because they neither can get infected by susceptible nor by vaccinated persons. But herd immunity still arises, because generally it emerges successively by an interplay of contacts and transmissions:

- The infected persons meet a certain number of persons. If there are vaccinated people in the population, there tend to be less susceptible persons among their contact persons than in a case without vaccinations. Thus, less transmissions happen.
- Consequently, susceptible people meet less infected people (because less people got infected) and so they have a lower probability to get infected themselves.

**An alternative interpretation of this effect:** The benefit of unvaccinated people is a consequence of the fact, that unvaccinated susceptible people meet vaccinated people. In a case without vaccination, these vaccinated people might have been infectious and could have infected a susceptible person.

## 4.4 Calculation of the herd immunity

### 4.4.1 Definition

Herd immunity should describe the benefit of the unvaccinated part of the population. In the context of epidemics, people only receive a disadvantage. Hence, the benefit can only be a decrease of this disadvantage. This decrease of a quantity always requires a comparison of different situations.

What the disadvantage is and how it can be calculated remains unresolved at the moment - this will be discussed later in chapter 4.4.2. The only requirement is the possibility to specify the disadvantage for each point in time of the system.

The proposed calculation of herd immunity needs three steps:

1. Two scenarios are required: a base scenario and a test scenario. The base scenario represents the initial situation and the test scenario is a situation with a vaccination strategy that might lead to a benefit.
2. Calculate the disadvantage of the unvaccinated population for both scenarios in each point in time.
3. Calculate the herd immunity as the relative decrease of the disadvantage from the basis scenario to the test scenario.

This leads to the following equation:

$H(t)$  . . . . Herd immunity in time  $t$   
 $T_0(t)$  . . . . Disadvantage of the unvaccinated population in the base scenario in time  $t$   
 $T_V(t)$  . . . . Disadvantage of the unvaccinated population in the test scenario in time  $t$

$$H(t) = \frac{T_0(t) - T_V(t)}{T_0(t)}$$

### Consequences of this definition

- Herd immunity is a time-dependent function which represents the effect of herd immunity for a given scenario in each point in time.
- This way of calculation is suitable for discrete-time models (like agent based models) and continuous-time models (like ODE-models) as well as for real studies with a control group.
- The function of herd immunity can be discrete-time or continuous-time, depending on the model, study or scenario.
- Herd immunity ranges over the interval  $(-\infty; 1]$ 
  - ◊  $= 1$  indicates a maximal effect for unvaccinated people. They do not have any disadvantage by disease any more.
  - ◊  $> 0$  indicates a positive effect for unvaccinated people. They experience a decrease of the disadvantage.
  - ◊  $= 0$  indicates no effect for unvaccinated people, their disadvantage does not change.
  - ◊  $< 0$  indicates a negative effect for the unvaccinated people. They experience a growth of the disadvantage.
- Often, the base scenario is a situation without vaccination and the test scenario is a situation with vaccinations. But, it is possible that in the base scenario people are vaccinated already and in the test scenario even more people get vaccinated. Then, herd immunity should describe the effect of these excess vaccinations. The definition above allows to consider such situations because it compares unvaccinated people in both scenarios.
- Complete herd immunity, which means an extinction of the disease, is reached if and only if the herd immunity is 1.

**Remark:** The herd immunity is defined on the unvaccinated population but potentially it seems possible to calculate it on other population parts as well. This could be the whole population for a generalized view or subgroups of the unvaccinated population to provide detailed aspects. It will be subject to further research to find out which calculations are most appropriate.



### 4.4.2 Quantification of the disadvantage

The quantification of the disadvantage that the unvaccinated population experiences is obviously one of the central aspects for the calculation of herd immunity.

**Statement:** The disadvantage is represented by persons that get infected with the disease. People that never get in touch with the disease cannot contribute to the disadvantage.

However, the assessment of an epidemic in a point in time is not necessarily bound to this single point. Instead, it might also depend on the past. Technically, there are three possibilities regarding consideration of infections over the time:

- **New infections:** The new infections refer to the number of persons that get infected in the concerned time point or time step (for discrete-time systems). In continuous, differentiable systems this corresponds to the derivative of the infections.
- **Infections since the beginning of the epidemic:** This calculation refers in a time point to the number of persons that have been infected since the start of the epidemic inclusively up to this point in time. This makes sense if the epidemic is considered as a completed event where the total number of infections are relevant.
- **Infections during the last  $x$  time units:** This calculation refers in a time point to the number of persons that have been infected in the last  $x$  time units including the actual point in time. If this method is used  $x$  should be chosen accordingly to the actual epidemic.

#### Usage of the different temporal considerations

- **New infections:** This method makes sense, if only the actual point in time is relevant for assessment of an epidemic.
- **Infections since the beginning of the epidemic:** This method makes sense, if the epidemic is considered as a completed event where the total number of infections are relevant.
- **Infections during the last  $x$  time units:** This consideration makes sense, if the epidemic underlies variations and the infection numbers are relevant over a time interval for assessment of the epidemic.

**Calculation of the disadvantage:** The calculation is done in two steps:

1. Choose the suitable temporal resolution at first.

2. Calculate the disadvantage as the proportion of the number of infections in the unvaccinated population based on unvaccinated persons:

$$T(t) = \frac{\text{Number of infections of unvaccinated persons within the chosen time interval}}{\text{Number of unvaccinated persons}}$$

**Remarks:**

- Especially for long time intervals, it can happen the  $T(t) > 1$  when people are infected multiple times because they get susceptible again after infection.
- The normalization of the disadvantage on the unvaccinated population - this is the calculation on the proportion based on unvaccinated people - is essential because the number of unvaccinated people in the base and the test scenario is usually different. Consequently, a comparison of absolute infection numbers is not correct.

### 4.4.3 Alternative calculation of the disadvantage

Often, the number of infections is unknown. Then it is still possible to use different quantities for the calculation of herd immunity. But one has to keep in mind that these calculation might produce misleading results. Below, some alternative quantities are discussed.

**Carrier rate:** The carrier rate describes the number of infected people within a population. Its change is composed of infections and recoveries. Herd immunity can be calculated by the carrier rate if the infections are not explicitly known. For that method all carriers have to be counted.

**Error:** People stay infected until they recover. For a uniform recovery time the function of the herd immunity is a shifted version of the original herd immunity by the recovery time. If recovery time is variable, the herd immunity is also distorted.

**New sickness cases:** Another option is quantification of the disadvantage by new sickness cases. One should keep in mind that not all infected people necessarily fall sick.

**Error:** If the proportion of infected people who fall sick is the same in both scenarios, then this consideration is a good approximation for herd immunity. However, the curve might be shifted and even distorted due to incubation time.

**Number of sick people:** Counting the number of sick people in a point in time is another alternative way to quantify the disadvantage.

**Error:** This calculation results in a herd immunity that is shifted and possibly distorted by both the incubation time and recovery time.

## 4.5 Herd immunity in models

### 4.5.1 Usage

The explanations show clearly that herd immunity is an effect which occurs because people interact with each other and transmissions happen upon these contacts. A model that simulates epidemics where herd immunity can occur has to take account of these relationships. Agent based models seem to fit perfectly because they represent persons naturally in their interactions and the consequences. Furthermore they are flexible, easily extendable and allow numerous analyses of details that can show the spread of an epidemic from every angle. Therefore, the studies in the following chapters are done by agent based models.

Models in this work serve as tools that allow to make epidemiologic effects, especially herd immunity, visible and to examine them.

### 4.5.2 Prerequisites

**Serotypes:** In this chapter we consider models with only one existing serotype. Hence, serotype shift cannot occur because a serotype cannot be replaced by another. Still, herd immunity is likely to appear in many situations.

**Impact of vaccinations:** The assumption for this model is that vaccinations either have full impact or no impact at all for a particular person. This means that a person with an immune response upon a vaccination cannot be infected any more. Other persons will still be considered to be susceptible.

**Vaccinated persons:** The term “vaccinated person” in this study refers to persons who produced an immune reaction after a vaccination so that they became resistant. Persons who received a vaccination that had no effect on them are not separately considered, they are just found among unvaccinated people in the model. Such a strategy is justified because vaccinations without effect do not affect the propagation of the disease at all. They play a role for economic evaluations and resources in the healthcare system but these aspects do not matter in this study.

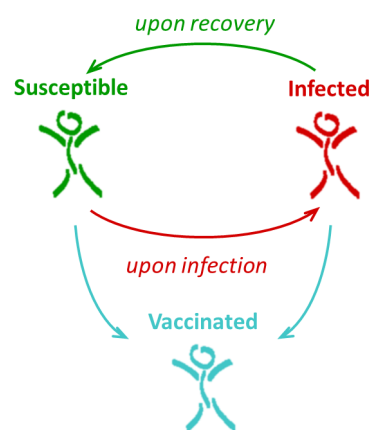
**Assumptions for modeling:** Consider a simple, agent based SIS or SIR model. This models works under the following assumptions:

- The agents in the model are persons.
- All persons in the model are of the same type. This means, they do not have different behavioral rules or sets of attributes.

- For SIS diseases, three infection states are available: *susceptible*, *infected* and *vaccinated*. For SIR epidemics, the state *resistant* additionally exists. Every person is in exactly one of those states in each point in time.
- Contacts are distributed uniformly and randomly on the population. The used parameter is  $k =$  “contacts per person per time unit”.
- The propagation of the disease happens upon three rules:
  - If a susceptible person meets an infected person, then the susceptible person changes its infection state to *infected* with probability  $\alpha$ .
  - An infected person stays exactly  $d$  time units in this state until it changes its state to *susceptible* (in SIS epidemics) or *resistant* (in SIR epidemics).
  - A vaccinated person stays in the state *vaccinated* until the end of the simulation.

Illustrations of the infection states of SIS and SIR models are shown in figures 4.2 and 4.3.

- The model is simulated in discrete time steps. Personal updates happen once in each time step after all contacts are performed. Contacts are considered to happen simultaneously for all persons, and so do personal updates.

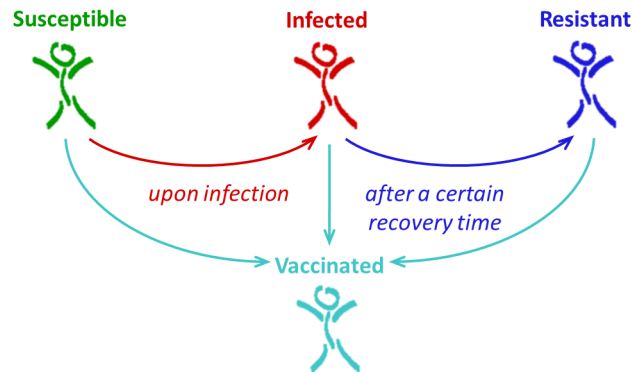


**Figure 4.2:** Infection states of persons in an SIS model

## 4.6 Herd immunity as an explaining quantity

### 4.6.1 Introduction

Based on the definitions in chapter 4.2, this chapter focuses on possibilities for calculations and concrete representations of herd immunity. The aimed benefit of this part of the study is to gain a deeper understanding of the system and on the spread of epidemics



**Figure 4.3:** Infection states of persons in an SIR model

by calculation of herd immunity. It should also help to understand more about the functionality and the dynamics of epidemic models. SIS and SIR epidemics will serve as examples because they are well known, extensively examined and easy to handle.

#### 4.6.2 Analytic calculation of herd immunity in an SIS model

**Previous knowledge:** It is known that an SIS model reaches an equilibrium after an adaption phase. The equilibrium is stable as long as there are no changes in the model. If persons are vaccinated during runtime, another adaption phase leads to a new equilibrium.

**Calculation of the equilibrium:** The following notation is used for the calculation:

- $N$  Number of persons
- $S$  Number of susceptible persons
- $I$  Number of infected persons
- $V$  Number of vaccinated persons
- $k$  Number of contacts per person per time unit
- $\alpha$  Infection probability
- $d$  Recovery time

The equilibrium is inevitably reached if and only if the numbers of newly infected persons and recovered persons are the same in each time step.

Statistically these equations hold for one time step:

$$\begin{aligned} \text{new infections} &= S \cdot k \cdot \frac{I}{N} \cdot \alpha \\ \text{recoveries} &= \frac{I}{d} \text{ with } d \geq 1 \end{aligned}$$

Thus, the equilibrium is reached when equation 4.1 is true.

$$S \cdot k \cdot \frac{I}{N} \cdot \alpha = \frac{I}{d} \tag{4.1}$$

Solving the equation for  $S$  leads to another equation 4.2. This only works if  $I > 0$ . The situation  $I = 0$  is trivial because in this case persons can never get infected. Then, every calculation is useless.

$$S = \frac{N}{d \cdot k \cdot \alpha} \text{ and } I > 0 \quad (4.2)$$

This formula leads to a few important statements about the model:

- $S$  cannot be greater than  $N$ . Also,  $d \cdot k \cdot \alpha \geq 1$  must be true, otherwise the equation cannot yield a result.  $d \cdot k \cdot \alpha < 1$  means that the epidemic would die out and  $S$  would equal  $N - V$ . Then,  $d \cdot k \cdot \alpha = 1$  applies.
- $S + I + V = N$  must always be true. Since  $N$  and  $V$  are assumed to be unchanged for a whole simulation it is possible to calculate the number of infected people in the equilibrium:  $I = N - S - V$ .  $I$  cannot be negative and so  $N \geq S + V$  must be true. The number of vaccinated people can be chosen freely from 0 to  $N$ . If  $N - V$  is lower than the calculated  $S$ , then  $I$  would be negative. This means that the disease becomes extinct.
- This leads to equation 4.3.

$$S = \begin{cases} \frac{N}{d \cdot k \cdot \alpha} & \text{if } d \cdot k \cdot \alpha > 1 \text{ and } V \leq N - \frac{N}{d \cdot k \cdot \alpha} \\ N - V & \text{otherwise} \end{cases} \quad (4.3)$$

- It is possible to show that the calculated equilibrium for  $S$  is independent on the initial state of the system. For this purpose show: If  $S$  is lower (respectively greater) than the calculated equilibrium, then more (respectively less) people get newly infected than people recover in this time step.

**Attention:** The following calculation only applies for the equilibrium and not for the adaption phase.

**Assumptions:**

- Require  $d \cdot k \cdot \alpha > 1$  and for the initial configuration  $I \neq 0$  so that  $S$  can reach the calculated equilibrium.
- **Base scenario:** No vaccinated persons.
- **Test scenario:** Vaccination of  $V$  persons ( $V \leq N$ ).

**Quantification of the disadvantage:**

- **Type:** As a comparison, the disadvantage is calculated upon both the new infections and the carrier rate.
- **Temporal resolution:** The disadvantage is calculated for a single point in time (and not over intervals).

**Notations for the state in the equilibrium:**

- $T_0$  Disadvantage of the unvaccinated population in the base scenario  
 $T_V$  Disadvantage of the unvaccinated population in the test scenario  
 $I_0$  Number of infected people in the base scenario  
 $I_V$  Number of infected people in the test scenario  
 $H$  Herd immunity

**Calculation of the herd immunity in the equilibrium:**

Disadvantage by new infections	Disadvantage by carrier rate
<p>The disadvantage is all new infections based on the unvaccinated population.</p> $T_0 = \frac{I_0}{N} = \frac{N - S}{d \cdot N}$ $T_V = \frac{I_V}{N - V} = \frac{N - S - V}{d \cdot (N - V)}$ $H = \frac{T_0 - T_V}{T_0}$ $= \frac{\frac{N-S}{d \cdot N} - \frac{N-S-V}{d \cdot (N-V)}}{\frac{N-S}{d \cdot N}}$ $= \frac{S \cdot V}{(N - S) \cdot (N - V)}$	<p>The disadvantage is the carrier rate based on the unvaccinated population.</p> $T_0 = \frac{I_0}{N} = \frac{N - S}{N}$ $T_V = \frac{I_V}{N - V} = \frac{N - S - V}{N - V}$ $H = \frac{T_0 - T_V}{T_0}$ $= \frac{\frac{N-S}{N} - \frac{N-S-V}{N-V}}{\frac{N-S}{N}}$ $= \frac{S \cdot V}{(N - S) \cdot (N - V)}$

We see that the herd immunity is the same with both calculation methods. This is because the time shift due to the recovery time does not affect the herd immunity in the equilibrium.

Calculation of the herd immunity with the formula is only possible if the disease does not become extinct. If  $V$  is so great that the disease becomes extinct then herd immunity equals 1. Hence, the following formula 4.4 for herd immunity in the equilibrium of SIS-Models applies (considering the calculation of  $S$  according to equation 4.3).

$$H = \begin{cases} \frac{S \cdot V}{(N-S) \cdot (N-V)} & \text{if } 0 \leq V \leq N - \frac{N}{d \cdot k \cdot \alpha} \\ 1 & \text{if } N - \frac{N}{d \cdot k \cdot \alpha} < V \leq N \end{cases} \quad (4.4)$$

**Remark 1:** Complete herd immunity, which refers to the situation when the disease becomes extinct, is reached when  $V > N - \frac{N}{d \cdot k \cdot \alpha}$ .

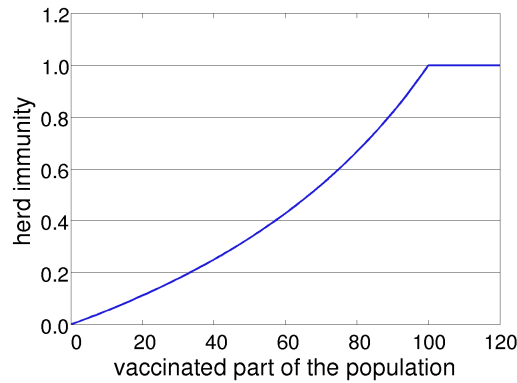
**Remark 2:** In the equilibrium  $N$  and  $S = \frac{N}{d \cdot k \cdot \alpha}$  have always the same value, independently of  $V$ . Hence the formula can be rearranged:

$$\begin{aligned} \frac{S \cdot V}{(N - S) \cdot (N - V)} &= \frac{\frac{N}{d \cdot k \cdot \alpha} \cdot V}{(N - \frac{N}{d \cdot k \cdot \alpha}) \cdot (N - V)} \\ &= \frac{N \cdot V}{(d \cdot k \cdot \alpha) \cdot N \cdot (1 - \frac{1}{d \cdot k \cdot \alpha}) \cdot (N - V)} \\ &= \frac{V}{(d \cdot k \cdot \alpha - 1) \cdot (N - V)} \end{aligned}$$

Furthermore herd immunity in the equilibrium of an SIS model can be seen as a function of  $V$  with the constants  $d$ ,  $k$  and  $\alpha$  with  $d \cdot k \cdot \alpha > 1$  (equation 4.5).

$$H(V) = \begin{cases} \frac{V}{(d \cdot k \cdot \alpha - 1) \cdot (N - V)} & \text{für } 0 \leq V \leq N - \frac{N}{d \cdot k \cdot \alpha} \\ 1 & \text{für } N - \frac{N}{d \cdot k \cdot \alpha} < V \leq N \end{cases} \quad (4.5)$$

**Remark 3:** The function of the herd immunity  $H(V)$  is a strictly monotonically increasing, convex function on  $0 \leq V \leq N - \frac{N}{d \cdot k \cdot \alpha}$  (see figure 4.4).



**Figure 4.4:** Graph of the herd immunity depending on the vaccinated part of the population. Chosen constants:  $d = 10$ ,  $k = 10$ ,  $\alpha = 0.02$

### 4.6.3 Example: Herd immunity in an agent based SIS model

**Assumptions:** In contrast to the theoretic calculations of chapter 4.6.2, in this chapter a concrete agent based model is used. The model builds on the description in chapter 3.8.1 and follows the specifications for SIS models in chapter 4.5.2.



**Simulation-parameters:**

Persons:  $N = 500\,000$

Contacts per person per time step:  $k = 10$

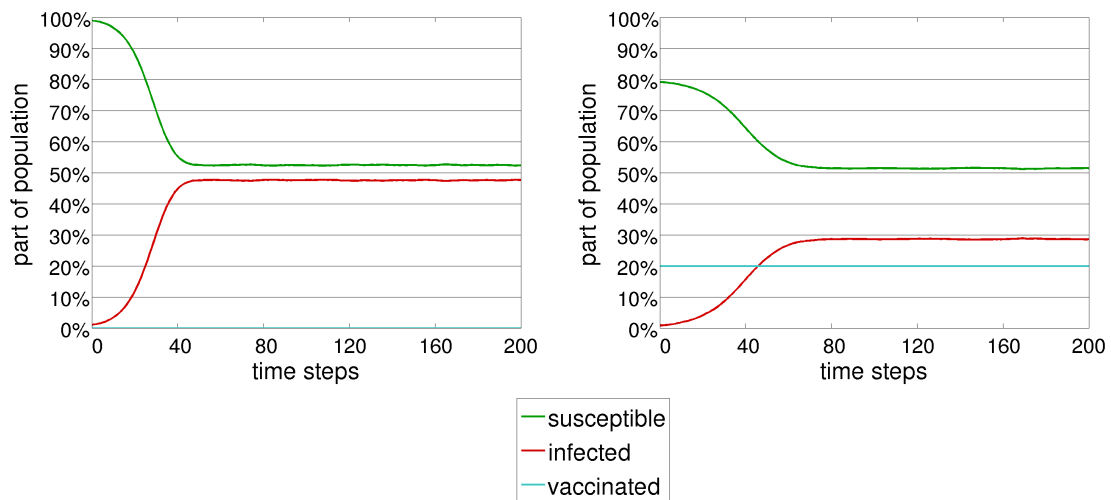
Infection probability:  $\alpha = 0.02$

Recovery time:  $d = 10$  time steps

Vaccinations: No vaccinations in the base scenario, 20% vaccinated people in the test scenario.

**Simulation results:** The results of the simulations are shown in figure 4.5. The graph of the states is a common visualization approach for epidemic models. This study focuses on the disadvantage caused by the disease. The previous chapter states that the disadvantage is generally represented by the number of infections and only alternatively by carriers. Hence, the number of infections per time step is plotted in figure 4.6.

In both scenarios the models reach their equilibrium after a short adaption phase. The figures show clearly that in case of vaccinations less infections happen and, consequently, the carrier rate is lower.



**Figure 4.5:** States in the SIS model  
**Left:** Base scenario (no vaccinated persons)  
**Right:** Test scenario (20% vaccinated persons)

**Definition of the disadvantage:** Three different methods for quantification of the disadvantage are used for comparison.

1. **Current new infections:**

The disadvantage in time  $x$  is  $\frac{\text{Number of newly infected persons in time step } x}{\text{Number of unvaccinated persons}}$

2. **New infections since start:**

The disadvantage in time  $x$  is  $\frac{\text{Number of newly infected persons since simulation start}}{\text{Number of unvaccinated persons}}$

**3. Current carriers:**

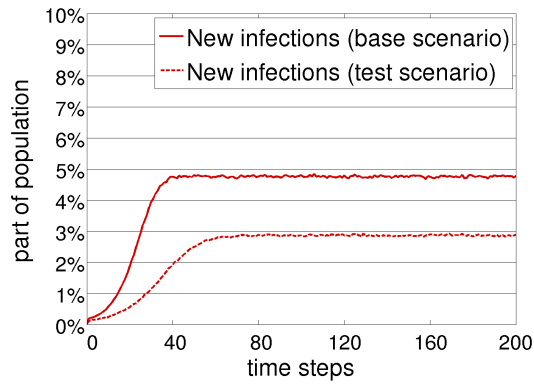
The disadvantage in time  $x$  is  $\frac{\text{Number of carriers in time step } x}{\text{Number of unvaccinated persons}}$

**Interpretation of the herd immunity:** Figures 4.7, 4.8 and 4.9 show the herd immunity based on the three different quantification methods.

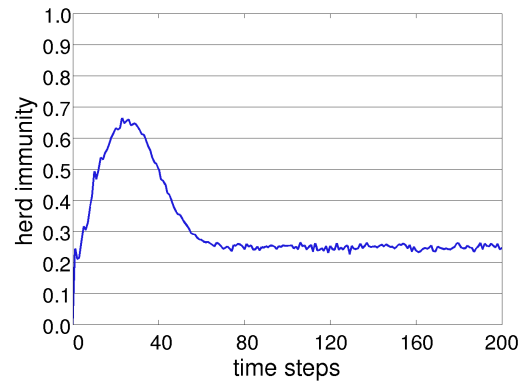
- All three cases show an increase of the herd immunity in the first few time steps up to 0.60-0.65 and then a decrease to about 0.25. This means that the disadvantage in the equilibrium (after about 70 time steps) for unvaccinated people is about 25% lower in the test scenario. In other words, the probability of an infection for unvaccinated people is 25% lower in case of 20% vaccination rate.
- The strong increase of herd immunity followed by a decrease can be explained easily because the adaption phase is shorter in the base scenario. Consequently, the advantage is greater during the adaption phase than in the equilibrium.
- Calculation of herd immunity by current new infections and carriers is very similar, only for calculation upon carriers the graph is slightly shifted to the right.
- If herd immunity is calculated upon all infections since simulation start, the decrease happens much slower because the high benefit in the beginning affects every later time step.

**Alternative Simulation:** Alternatively, the dynamic of herd immunity in SIS models might be shown by vaccinations during runtime, after the system has reached its equilibrium. Then it is possible to observe how the epidemic goes from one equilibrium to another under the impact of herd immunity. The simulation parameters are the same as above and the already shown simulation without vaccinations serve as the base scenario. Figure 4.10 shows the states and figure 4.11 shows the new infections of this simulation. The results make clear that in the time steps after the vaccination, a process occurs that finally brings the system successively to the new equilibrium. This circumstance can be seen explained by the herd immunity which is 0.2 directly after the vaccination and then rises to 0.25 (figure 4.12).

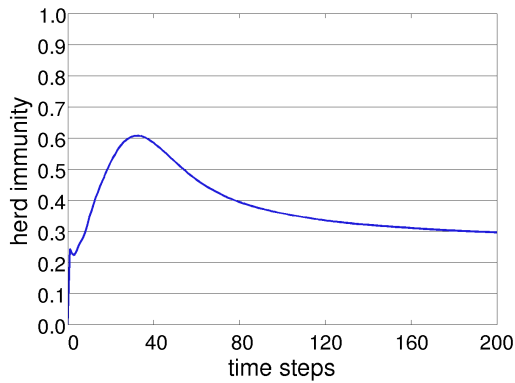
**Comparison with the analytic considerations:** The analytic calculation in chapter 4.6.2 provides for herd immunity of an SIS model in the equilibrium the formula  $H = \frac{V}{(d \cdot k \cdot \alpha - 1) \cdot (N - V)}$ . With  $d = 10$ ,  $k = 10$ ,  $\alpha = 0.02$  and a vaccination rate of 20% this yields the herd immunity  $H = \frac{100000}{(10 \cdot 10 \cdot 0.02 - 1) \cdot (400000)} = 0.25$ . This agrees very well with the results of the agent based model.



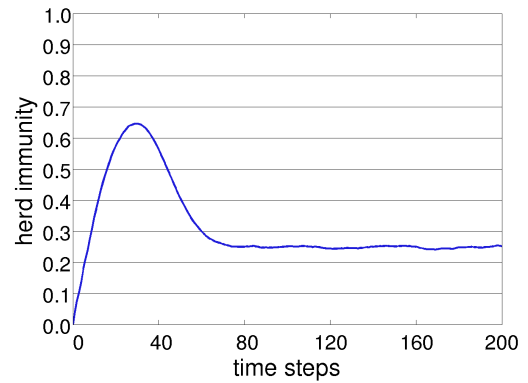
**Figure 4.6:** New infections in the SIS model



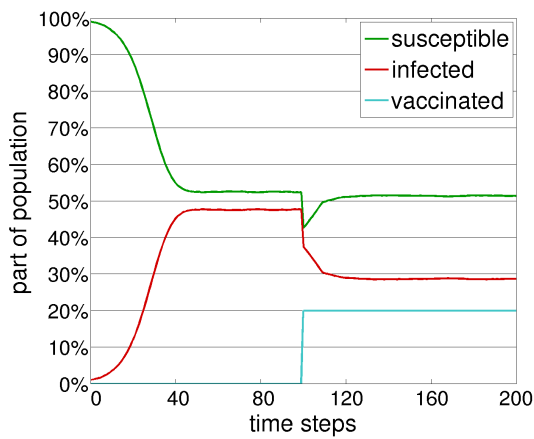
**Figure 4.7:** Calculation of the herd immunity upon current new infections



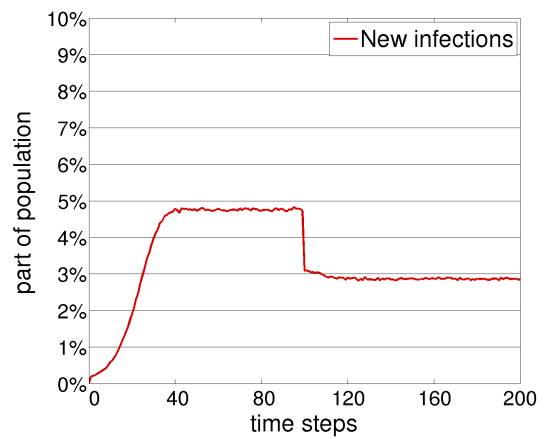
**Figure 4.8:** Calculation of the herd immunity upon infections since simulation start



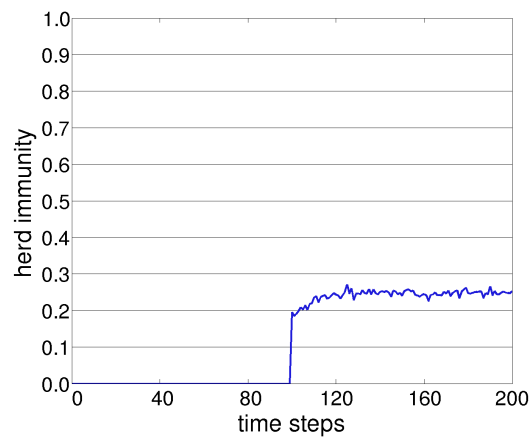
**Figure 4.9:** Calculation of the herd immunity upon current carriers



**Figure 4.10:** States in the SIS model with vaccinations after 100 time steps



**Figure 4.11:** Infections in the SIS model with vaccinations after 100 time steps



**Figure 4.12:** Herd immunity based on current infections with vaccinations after 100 time steps

#### 4.6.4 Example: Herd immunity in an agent based SIR model

**Assumptions:** Consider the simple agent based SIR model, which is defined in chapter 4.5.2.

**Simulation parameter:**

Persons:  $N = 500\,000$

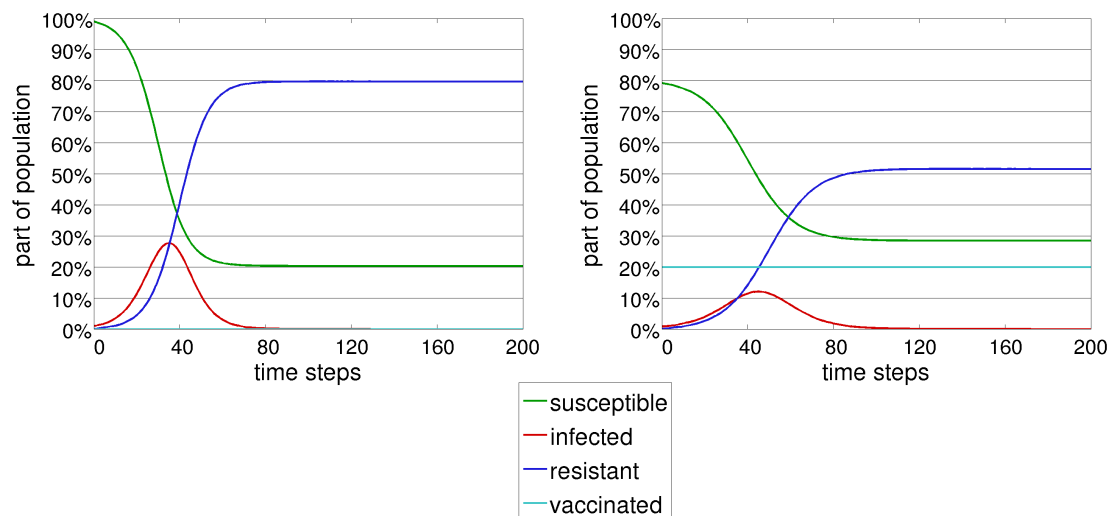
Contacts per persons per time step:  $k = 10$

Infection probability:  $\alpha = 0.02$

Recovery time:  $d = 10$  time steps

Vaccinations: No vaccinations in the base scenario, 20% vaccinated people in the test scenario.

**Simulation results:** The graphs in figures 4.13 show the results of the simulations by evaluation of the states. Figure 4.14 shows the infections per time step to visualize the disadvantage for the population. Both scenarios yield the well known curves of SIR epidemics. In case of vaccinations the peak of the epidemic is much lower but the epidemic lasts longer until it becomes extinct.



**Figure 4.13:** States in the SIR model

**Left:** Base scenario (no vaccinated persons)

**Right:** Test scenario (20% vaccinated persons)

**Remark:** The results clearly show that the epidemic lasts longer in case of vaccinations. Then, especially between time 40 and 80 more infections happen. This circumstance is shown by the blue area figure 4.15. So the main question is: “*Is there still a benefit by the vaccinations in time 40 to 80 or not?*” On the one hand there is a benefit because in the first 80 time steps much less people get infected in the test scenario, but on the other hand there are more infections in time 40 to 80. However, the difference is small in absolute numbers and large in relative values.

**Definition of the disadvantage:** Again, for comparison reasons three different methods are used to calculate the disadvantage.

1. **Current new infections**
2. **New infections since start**
3. **Current carriers**

**Interpretation of the herd immunity:** Figures 4.16, 4.17 and 4.18 show the results of the three calculations of herd immunity.

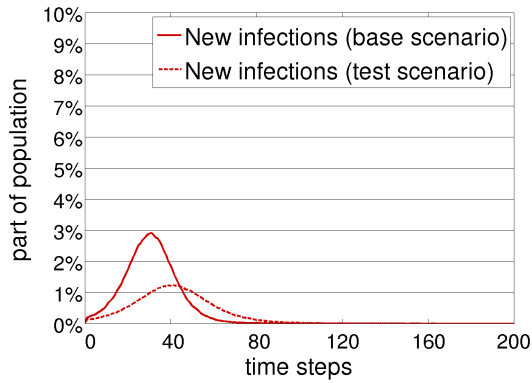
- The calculations upon current new infections or carriers are hardly meaningful. As expected, both graphs are very similar, and only the herd immunity based on carriers is slightly distorted. The less meaningful values of the herd immunity are based on the fact that the epidemic curve is shifted to the right by vaccinations.
  - In the first 40 to 50 time steps, most persons are infected and the herd immunity makes sense and is positive.
  - Between time 40 and 70, already more infections happen in the test scenario, hence the herd immunity is negative.
  - From time 80 onwards only very few infections happen and small number have to be compared. Still more infections happen in the test scenario, even though they are almost neglectable in absolute numbers. Hence, the herd immunity has values that are very negative and in case of  $\frac{x}{0}$  they are even virtually  $-\infty$ .
- Against this, the method with all infections since start clearly shows the benefit for the unvaccinated population. It is about 0.2 in the end. This means that in the case of vaccinations about 20% less infections happen to the unvaccinated people within the whole epidemic.

## 4.7 Conclusions

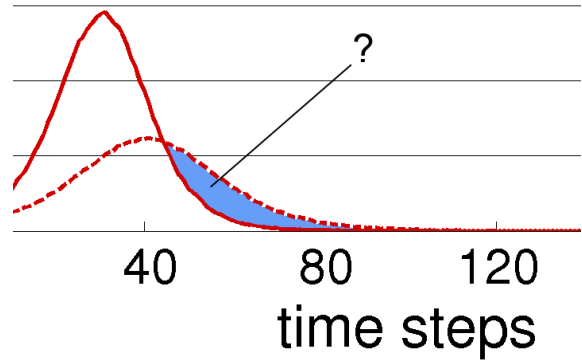
The examples obviously show an impact of vaccinations on the unvaccinated population. An appropriate calculation and representation of this impact by herd immunity highly depends in the actual epidemic and on the point of view. Especially temporal shifts, which apply for example for SIR epidemics, require a detailed definition of the interval that is used to represent the disadvantage at a point in time.

Generally, it is suggested to calculate the disadvantage upon newly infected people but if this information is not available other methods exist as alternatives. One should consider that these alternatives might yield different results than the original herd immunity.

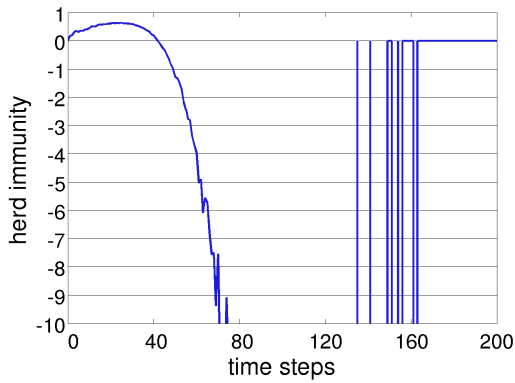
However, the examples reveal the capability of herd immunity and its additional information on the system if the temporal component is chosen appropriately. This leads to the main conclusions of this chapter:



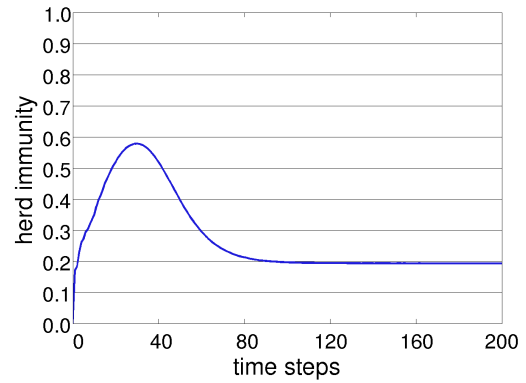
**Figure 4.14:** New infections in the SIR model



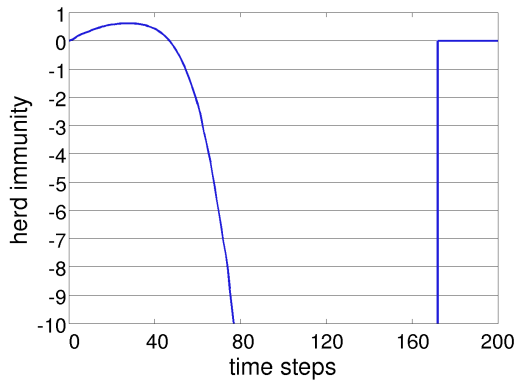
**Figure 4.15:** Is there a disadvantage or still a benefit in time 40 to 80?



**Figure 4.16:** Calculation of the herd immunity upon current new infections



**Figure 4.17:** Calculation of the herd immunity upon infections since simulation start



**Figure 4.18:** Calculation of the herd immunity upon current carriers

- **The temporal component must be chosen accordingly to the actual epidemic and given problem.**
- **Alternative quantifications of disadvantage than by newly infected people are possible but might lead to different results.**
- **Usage of an appropriate time interval for quantification of the disadvantage results in informative values that represent the benefit of unvaccinated people by a vaccination strategy.**

## 4.8 Parameter studies on herd immunity

### 4.8.1 Relationships in the SIS model

In this chapter, SIS epidemics in an agent based model are compared to the analytic formula for deeper understanding of the results in the following chapters. Formula 4.4 from the analytic calculation implies that the equilibrium of SIS models is a function of the product of infection probability, recovery time and contacts per person. But, simulations with the agent based model partly lead to different results than the formula. These differences will be examined and explained by two test series below.

#### **Test series 1 - Relationship between infection probability and contact number:**

A constant product of infection probability and contact number yields almost the same carrier rates by simulation with the agent based model, corresponding to the analytic formula. This behavior can easily be explained because the number of infections per time step is exactly the product of potential infectious contacts and the infection probability. Potentially infectious contacts are contacts between infected and susceptible persons. Obviously the probability of such a contact is  $2 \cdot \text{carrier rate} \cdot (1 - \text{carrier rate})$ . Thus the transmissions on the agent based model in one time step for a given carrier rate depend on the product of infection probability and contact number. Thus, the model should result in the same equilibrium.

Used settings:

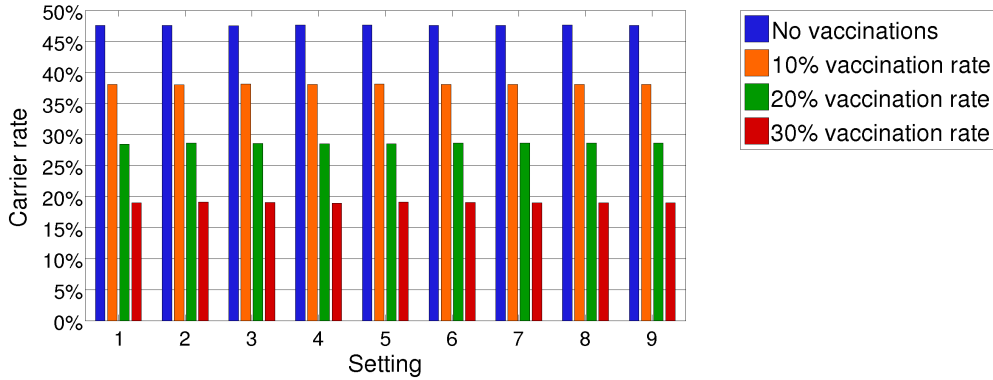
- Number of persons: 500 000
- Recovery time: 10 time steps
- The carrier rate is set accordingly to the analytic calculation to prevent a long adaption phase.
- Simulation time: 150 time steps
- Calculation of the carrier rate: Average carrier rate of the time steps 126 to 150.



Infection probability and contact number are varied in nine different settings (see table 4.1) and simulated for 0%, 10%, 20% and 30% vaccination rates with the agent based model.  $k$  defines the average number of contacts per person which is used to calculate the total number of contacts in the population, hence it can be a decimal number and does not need to be an integer. The results, which are represented by the carrier rates in the equilibrium, are shown in figure 4.19.

**Table 4.1:** Parameter settings of the test series for infection probability  $\alpha$  and contacts per persons  $k$ .

Setting	1	2	3	4	5	6	7	8	9
$k$	2	4	6	8	10	$\frac{100}{8}$	$\frac{100}{6}$	$\frac{100}{4}$	$\frac{100}{2}$
$\alpha$	$\frac{0.2}{2}$	$\frac{0.2}{4}$	$\frac{0.2}{6}$	$\frac{0.2}{8}$	0.02	0.016	0.012	0.008	0.004

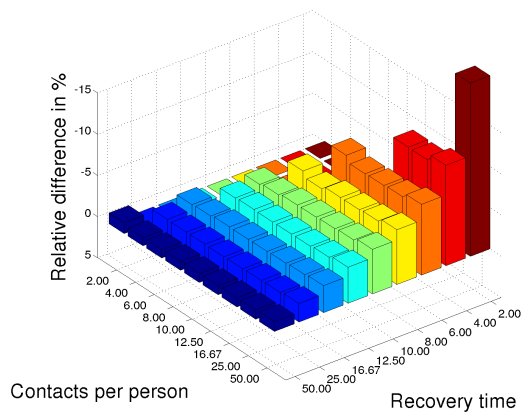


**Figure 4.19:** Results of the test series for infection probability and contacts per persons.

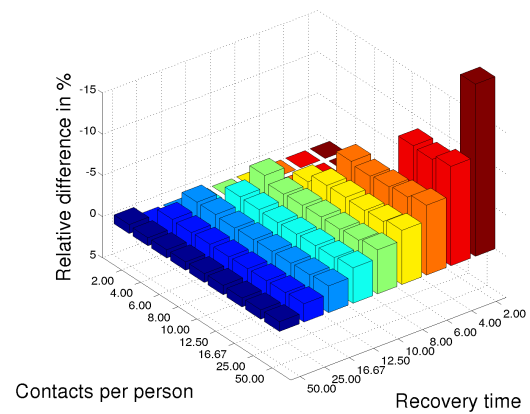
**Test series 2 - Study on the deviations of the equilibrium in the agent based model and the analytic calculation:** The deviations of the agent based model compared to the analytic calculation are examined in the second test series. Since the relationship between infection probability and contact number is known, it is enough to vary the recovery time and the contact numbers and leave the infection probability constant.

The general settings for the test series are the same as for the test series above. The parameter variation consists of all combinations of 2, 4, 6, 8, 10,  $\frac{100}{8}$ ,  $\frac{100}{6}$ ,  $\frac{100}{4}$  and  $\frac{100}{2}$  for both contacts per person and recovery time. If the recovery time is a decimal number then people receive in case of an infection the rounded down or up value, so that in average among the population the given number is reached.

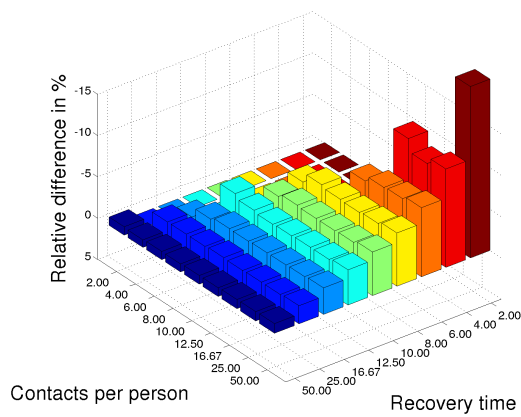
The result is shown as the relative difference of the carrier rate in the equilibrium of the agent based model from the analytic calculation. The whole test series is simulated for 0% (figure 4.20), 10% (figure 4.21), 20% (figure 4.22) and 30% (figure 4.23) vaccination rates. The z-axis is printed upside-down since the carrier rates in the agent based model are generally lower than the ones in the analytic calculation.



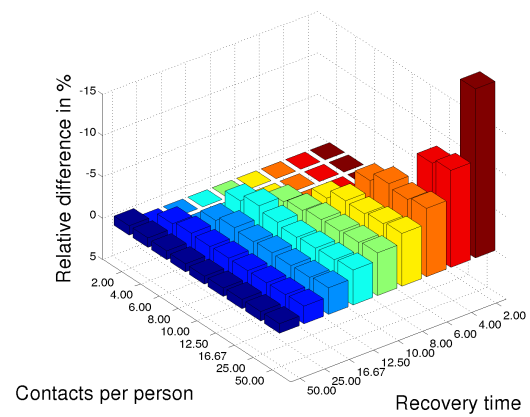
**Figure 4.20:** Deviation of the agent based model from the analytic calculation in case of no vaccinations.



**Figure 4.21:** Deviation of the agent based model from the analytic calculation in case of 10% vaccination rate.



**Figure 4.22:** Deviation of the agent based model from the analytic calculation in case of 20% vaccination rate.



**Figure 4.23:** Deviation of the agent based model from the analytic calculation in case of 30% vaccination rate.

**Interpretation of the results of the test series:** A few general insights on differences between the agent based model and the statistical calculation turn out.

- If the product of infection probability and contact number stays constant, then the agent based model yield the same results. This relationship can be explained very well.
- The carrier rate of the agent based model is generally lower than the one in the analytic calculation.
- Simulations without any difference are usually situations where the disease becomes extinct.
- The longer the recovery time is the greater is the difference.
- The difference is obviously independent from the contact number but strongly depends on the recovery time.
- The vaccination rate does not seem to have an impact on the difference.

**Explanation of the results:** The results of test series 1 are already explained by the description of this test series. The assumed relationship between infection probability and contact number is confirmed by the simulations.

It is expected that the carrier rate remains unchanged if the product of recovery time and contact number stays constant, because such a parameter variation corresponds to a distortion of the time axis. Multiplication of the contact number with  $x$  generates in one time step as many contacts as before in  $x$  time steps. Division of the recovery by  $x$  finally leads to a situation where one time step corresponds to  $x$  time steps in the original situation. Scaling of the time axis theoretically does not have an impact on the carrier rate. But the agent based model shows differences that are as larger as longer the recovery time is. A shorter recovery time corresponds to longer time steps in the real system, thus a rougher discretization and hence a larger error. Consequently, the smallest error happens for long recovery rates which represent a fine discretization. This leads to the conclusion that the differences of the carrier rates in both models are an effect of discretization of the agent based model.

### 4.8.2 Parameter studies with the SIS model

**Preliminary remarks:** It is known that SIS epidemics always reach an equilibrium. This quantity is more important for characterization of the epidemic than possibly existing adaption phases. Consequently, examinations on longer time intervals are less important because such an epidemic has no real start and end, in contrast to SIR epidemics. The numbers of new infections per time step in the equilibrium are meaningful results for

comparison of SIS simulations. Additionally they are the main characterization of the disadvantage caused by the disease.

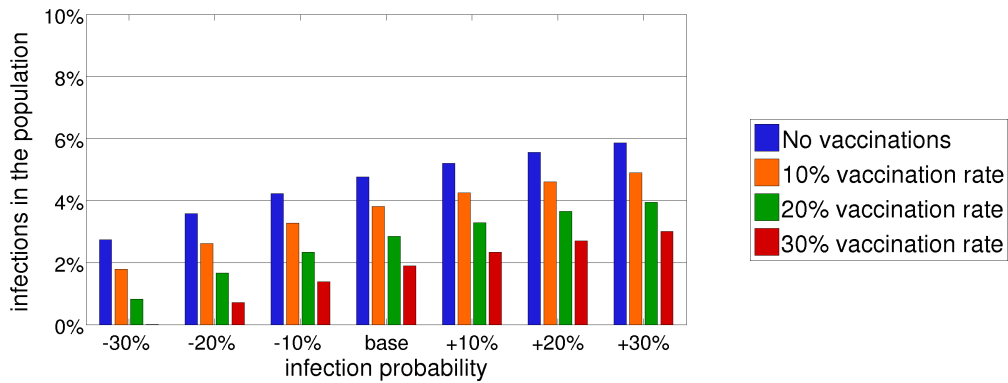
Herd immunity is suitable for comparison of effects of different vaccination rates or strategies. The preceding chapters have shown that all calculation methods for herd immunity of an SIS model yield useable results. However, herd immunity in this study is calculated upon current new infections to take the pure outcomes and because they allow statements about single points in time in the equilibrium.

The relationship between infection probability and contact numbers is already shown and explained. Thus, it is sufficient to vary only one of these variables and keep the other one constant. Hence, in this study only the infection probability and the recovery time are varied.

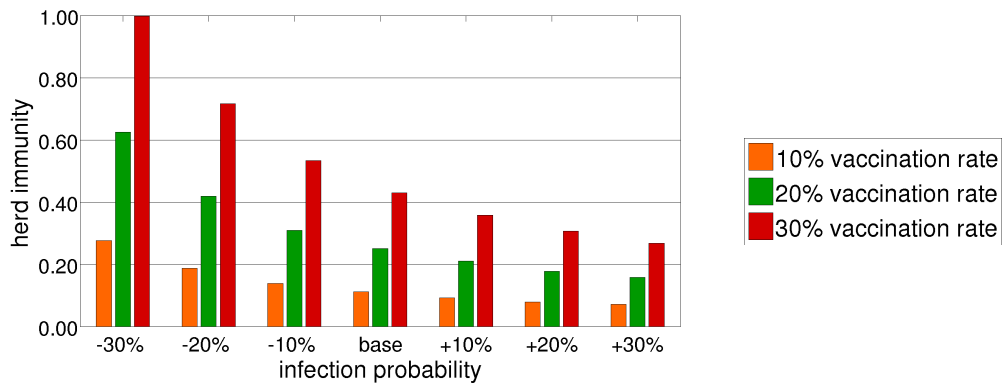
### Overview

- **Simulation:** SIS epidemic
- **Model:** Agent based model
- **Representation of the results:** New infections per time step in the equilibrium. Since random noise can occur for agent based models, the new infections are calculated by the mean of 25 time steps.
- **General settings:**
  - Persons: 1 000 000
  - Infection probability: 0.02
  - Contacts per person per time step: 10
  - Recovery time: 10 time steps
  - Vaccinations: All simulations are performed without vaccinations and 10%, 20% and 30% vaccination rate.
- **Herd immunity:**
  - Base scenario: Simulation without vaccinations
  - Test scenario: Simulations with vaccinations
  - Calculation method: Current new infections in the equilibrium.
- **Variations:** Both, infection probability and recovery time are varied by 10%, 20% and 30% lower and higher values. Combinations of these parameters are not varied.

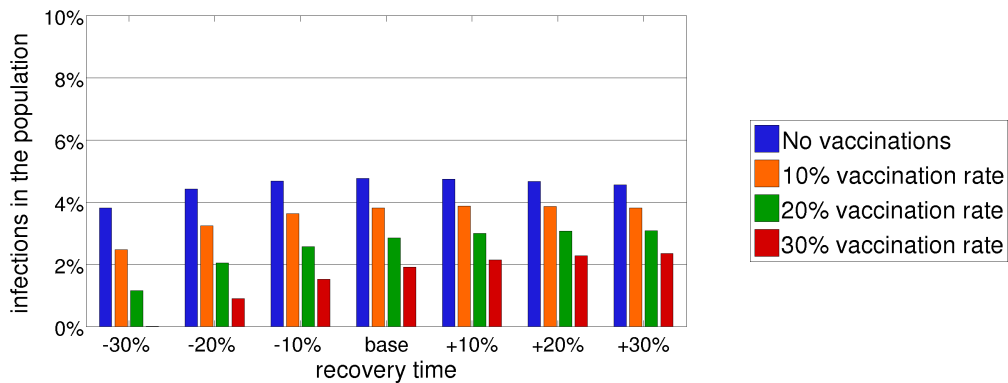
#### 4.8. PARAMETER STUDIES ON HERD IMMUNITY



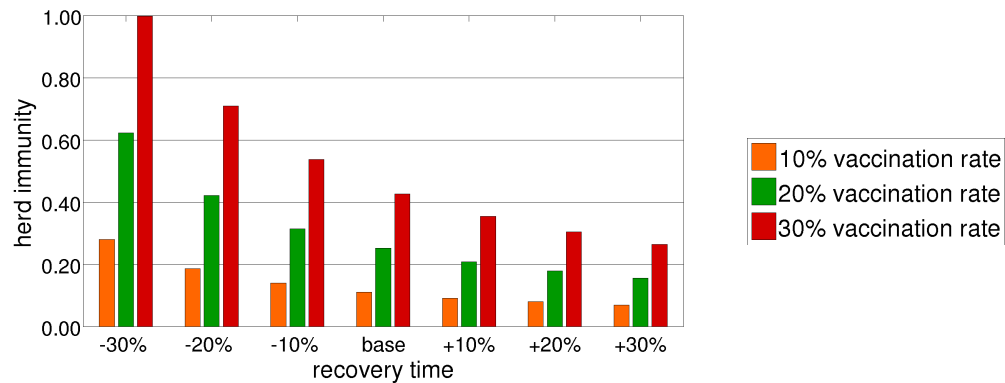
**Figure 4.24:** Variation of the infection probability in the SIS model. The bars show the new infections per time step in the equilibrium.



**Figure 4.25:** Herd immunity of the variation of the infection probability in the SIS model. The bars show the herd immunity upon current new infections in the equilibrium.



**Figure 4.26:** Variation of the recovery time in the SIS model. The bars show the new infections per time step in the equilibrium.



**Figure 4.27:** Herd immunity of the variation of the recovery time in the SIS model. The bars show the herd immunity upon current new infections in the equilibrium.

**Results:** Figure 4.24 shows the results of the parameter variation of the infection probability while figure 4.26 shows the ones of the variation of the recovery time. The herd immunity is plotted in figure 4.25 and figure 4.27. This test series provides a few new insights:

- The number of infections is as higher as higher the infection probability is.
- A higher recovery time leads, if at all, only to a slightly increased infection number. However, due to the longer time to recovery the carrier rate increases significantly at the same number of new infections.
- The herd immunity is a higher as higher the vaccination rate is.
- The herd immunity for a given vaccination rate is as higher as less carriers are present in the base scenario. This means, the benefit for the unvaccinated population is higher in SIS epidemics with lower carrier rates.
- Complete herd immunity can be observed in two cases where the epidemic becomes extinct at 30% vaccination rate. In these cases the herd immunity equals 1 which confirms previous deliberations.

**Comparison of the results with the theory:** The theory provides a formula to calculate the herd immunity in SIS epidemics (see chapter 4.6.2, formulas 4.4 and 4.5). The tables 4.2 and 4.3 show the calculated herd immunity compared to the measured value in the agent based model. The results agree very well except for small deviations which can be explained by random noise.

**Table 4.2:** Comparison of the herd immunity in the SIS model with the formula (parameter studies of the infection probability)

Infection probability	Vaccination rate	Herd immunity		Deviation
		AB model	Formula	
-30%	10%	0.28	0.28	+0.73%
	20%	0.63	0.63	-0.03%
	30%	1.00	1.00	+0.00%
-20%	10%	0.19	0.19	-1.13%
	20%	0.42	0.42	-0.39%
	30%	0.72	0.71	-0.28%
-10%	10%	0.14	0.14	+0.04%
	20%	0.31	0.31	+1.02%
	30%	0.53	0.54	+0.48%
basis	10%	0.11	0.11	-0.22%
	20%	0.25	0.25	-0.45%
	30%	0.43	0.43	-0.20%
+10%	10%	0.09	0.09	+0.62%
	20%	0.21	0.21	-0.78%
	30%	0.36	0.36	-0.39%
+20%	10%	0.08	0.08	+1.61%
	20%	0.18	0.18	+0.74%
	30%	0.31	0.31	+0.01%
+30%	10%	0.07	0.07	-2.49%
	20%	0.16	0.16	-0.74%
	30%	0.27	0.27	-0.15%

**Table 4.3:** Comparison of the herd immunity in the SIS model with the formula (parameter studies of the recovery time)

Recovery time	Vaccination rate	Herd immunity		Deviation
		AB model	Formula	
-30%	10%	0.28	0.28	-0.97%
	20%	0.62	0.63	+0.34%
	30%	1.00	1.00	+0.00%
-20%	10%	0.19	0.19	-0.60%
	20%	0.42	0.42	-1.14%
	30%	0.71	0.71	+0.65%
-10%	10%	0.14	0.14	-0.52%
	20%	0.31	0.31	-0.49%
	30%	0.54	0.54	-0.32%
basis	10%	0.11	0.11	+0.58%
	20%	0.25	0.25	-0.63%
	30%	0.43	0.43	+0.65%
+10%	10%	0.09	0.09	+1.44%
	20%	0.21	0.21	-0.08%
	30%	0.35	0.36	+0.91%
+20%	10%	0.08	0.08	+0.04%
	20%	0.18	0.18	+0.21%
	30%	0.30	0.31	+0.51%
+30%	10%	0.07	0.07	+0.73%
	20%	0.16	0.16	+0.79%
	30%	0.26	0.27	+1.34%



### 4.8.3 Parameter studies with the SIR model

**Preliminary remarks:** SIR epidemics are always temporal limited processes because at the certain point the epidemic becomes extinct. Traditionally, only a few infected people exist in the beginning, then the number of infected people grows until it reaches a peak and then it decreases successively to zero. For characterization of SIR-Epidemics, both values of the temporal propagation (peak and duration) as well as the total number of infections of the whole epidemic are useful. In this study, the total number of infections are used to compare simulations of the parameter studies.

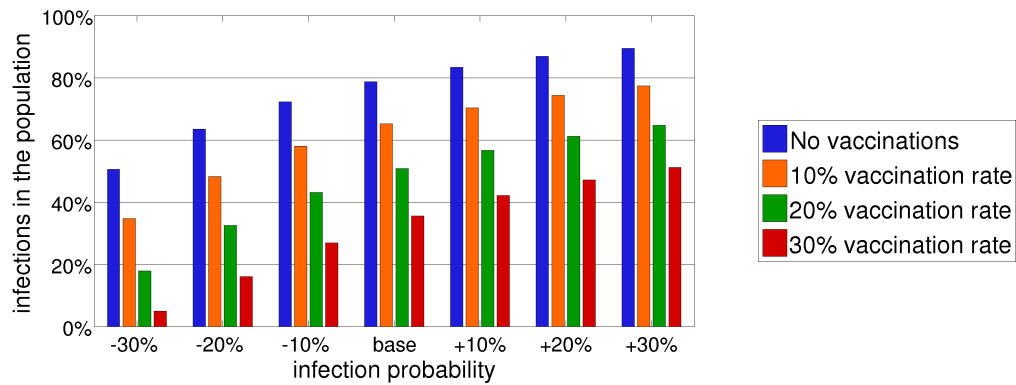
The calculation of herd immunity can be tricky for SIR models if one considers only the state in a single point in time because vaccinations can shift these kind of epidemics. The method to calculate herd immunity upon the total number of infections since start seems to be most promising since the SIR epidemic is always a time limited event. For comparison of different vaccination rates in the studies, the herd immunity at the end of the whole epidemic is used.

#### Overview

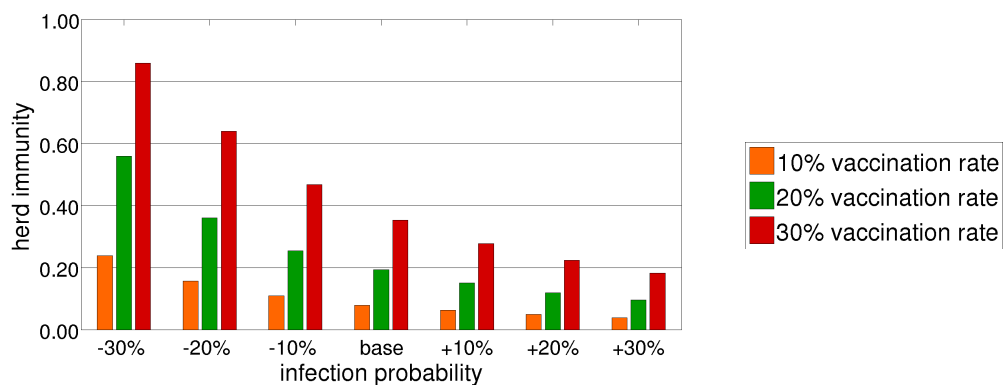
- **Simulation:** SIR epidemic
- **Model:** Agent based model
- **Representation of the results:** Total number of new infections for the whole epidemic.
- **General settings and variations:** The settings and parameter variations are the same as for the SIS model (see chapter 4.8.2). The start values are always 1% infected and 99% susceptible persons (vaccinations are not part of the start values, instead people get randomly vaccinated before the first time step).
- **Herd immunity:**
  - Base scenario: Simulation without vaccinations
  - Test scenario: Simulations with vaccinations
  - Calculation method: New infections since the start of the epidemic. For comparison only the herd immunity at the end of the epidemic is shown.

**Results:** Figure 4.28 shows the results of the variation of the infection probability while figure 4.30 shows those of the recovery time variation. The corresponding calculations of herd immunity are illustrated in figure 4.29 and 4.31. This test series provides new insights about SIR epidemics:

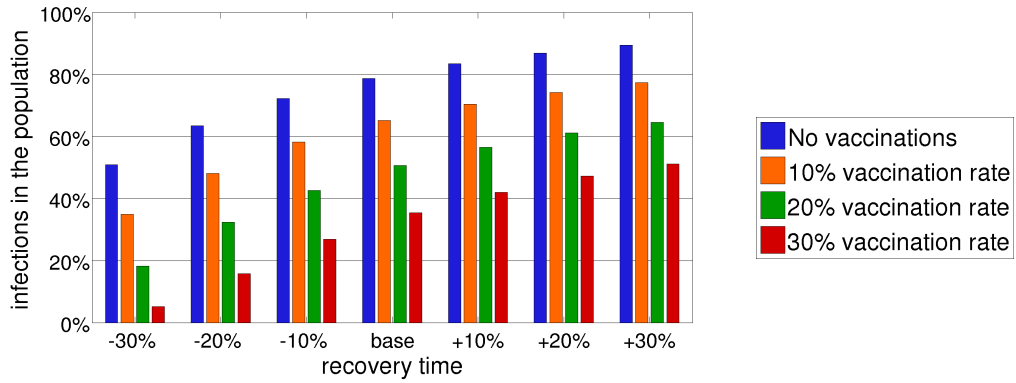
- Higher infection probability and recovery time leads to more total infected people during the epidemic.



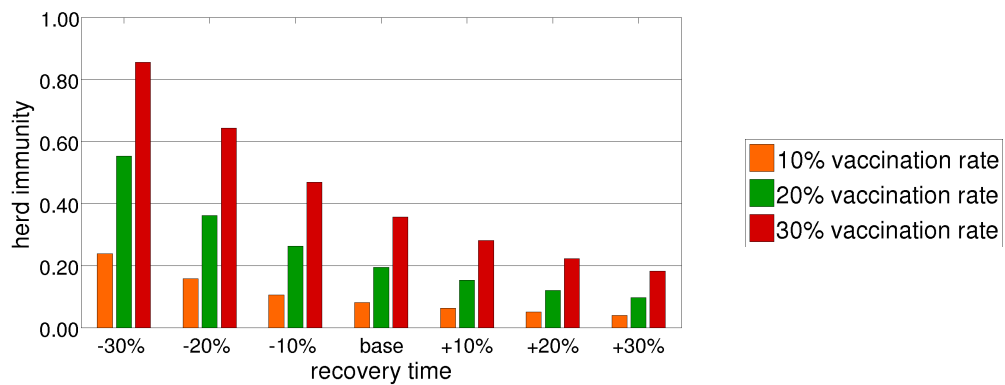
**Figure 4.28:** Variation of the infection probability in the SIR model. The bars show the total infections of the whole epidemic.



**Figure 4.29:** Herd immunity of the variation of the infection probability in the SIR model. The bars show the herd immunity upon new infections since start at the end of the epidemic.



**Figure 4.30:** Variation of the recovery time in the SIR model. The bars show the total infections of the whole epidemic.



**Figure 4.31:** Herd immunity of the variation of the recovery time in the SIR model. The bars show the herd immunity upon new infections since start at the end of the epidemic.

## CHAPTER 4. HERD IMMUNITY: TERMINOLOGY, MOTIVATION AND ORIGIN

---

- The herd immunity is as stronger as less infected people exist for the whole epidemic in the base scenario. This means, unvaccinated people have a lower benefit for fixed vaccination rates in epidemics with many infected people compared to ones with less infected people.
- The herd immunity is as stronger as higher the vaccination rate is.

# Chapter 5

## Comparison of Markov models and agent based models for epidemic simulations

### 5.1 Overview

This chapter provides a comparison of Markov models with the agent based approach for epidemic disease modeling, based on concrete examples and an analytic consideration. The infections in the agent based approach are a result of contacts between single persons and infection probabilities. In contrast to that, the Markov model requires direct parameterization of the probability that persons change from one state to another.

The comparison is done by the following procedure:

1. Construction of an agent based and a Markov model for simulation of an epidemic.
2. Parameterization of both models so that their results correspond to each other for simulations without vaccinations.
3. Implementation of different vaccination strategies on the population in both models and comparison of the results.
4. Presentation the different results and attempt to explain the differences from a mathematical point of view.

The goals of this study are firstly, to learn about the differences between contact based and non-contact based epidemic models and secondly, to understand the impact of herd immunity on the spread of epidemics better. Three concrete examples are studied for that reason:

1. Comparison of equilibriums in an SIS model.

2. Comparison of the temporal propagation of an SIR model.
3. Comparison of infection numbers of a complete season in an SIR model.

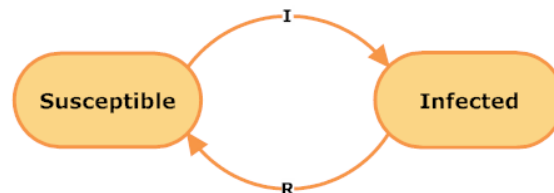
## 5.2 Comparison of equilibriums in an SIS epidemic

### 5.2.1 Construction of the models

**The agent based model:** The agent based model is the same that has been defined and constructed in chapter 4.5.2 and extensively examined in chapter 4.8.2. Hence, its functionality is not repeated again here.

**The Markov model:** The population is split into unvaccinated and vaccinated people that are simulated with different Markov models.

**Unvaccinated people:** The SIS Markov model consists of two states:  $s_1 = \text{“Susceptible”}$  and  $s_2 = \text{“Infected”}$ . Figure 5.1 shows the state transition diagram.



**Figure 5.1:** The state transition diagram of the SIS Markov model for unvaccinated people

The distribution of the system in a point in time  $t$  is mathematically represented by a 2-dimensional vector which contains the amounts of susceptible and infected people:  $\underline{a}_t = (a_{1,t}, a_{2,t})$ . The transition matrix contains the probabilities  $I$  that susceptible people becomes infected and  $R$  that infected people recover and become susceptible. Thus, the probabilities that people remain in their state are  $(1 - I)$  respectively  $(1 - R)$ .

$$M = \begin{pmatrix} 1 - I & I \\ R & 1 - R \end{pmatrix}$$

**Vaccinated people:** Vaccinated people always stay vaccinated, hence the Markov model consists of only state:  $s_1 = \text{“Vaccinated”}$ .

Figure 5.2 shows the tree of whole model.

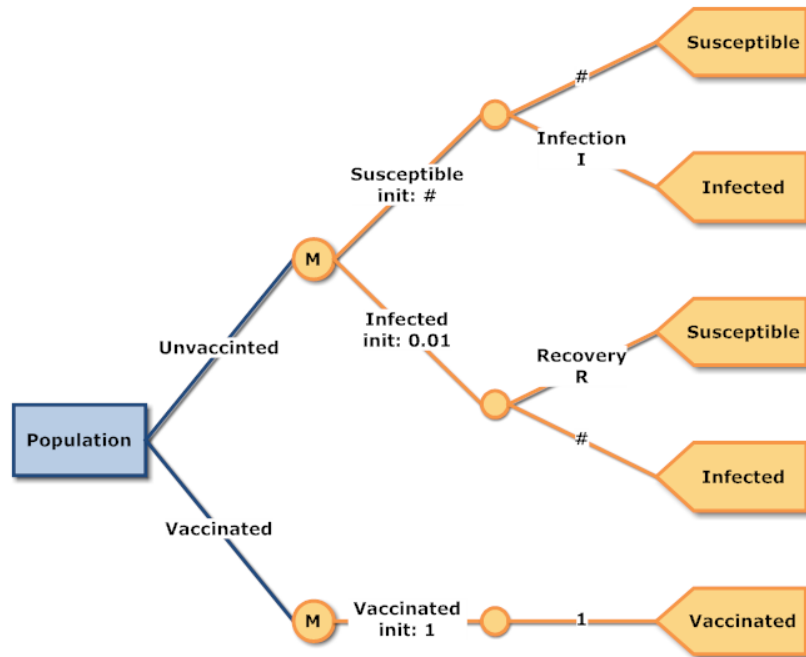


Figure 5.2: The Markov tree of the SIS Markov model

### 5.2.2 Parameterization of the model

**Parameterization of the agent based model:** The used parameters for this example are the same as in the base simulation in the last chapter:

- Contacts per person per time step:  $k = 10$
- Infection probability:  $\alpha = 0.02$
- Recovery time:  $d = 10$  time steps
- Vaccination rate: 0%
- Infected persons at simulation start: 1%  
However, start values are not important since only the equilibrium is relevant in this example.

**Parameterization of the Markov model:** The infection probability  $I$  and the recovery probability  $R$  in the Markov model must be set correctly so that the distribution reaches the same equilibrium as in the agent based model. Adaption phases to other equilibriums are not considered (they usually occur in the beginning of the simulation and after interventions during simulation).

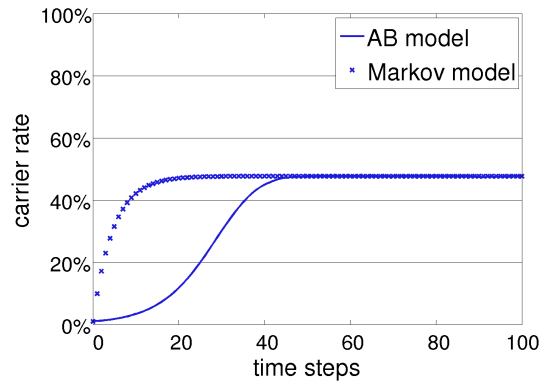
Thus, equation 5.1 defines for the agent based model two time-dependent functions  $I(t)$  and  $R(t)$ .

$$I(t) = \frac{\text{New infections}(t)}{\text{Susceptible people}(t-1)} \quad (5.1)$$

$$R(t) = \frac{\text{Recoveries}(t)}{\text{Infected people}(t-1)}$$

$I(t)$  represents the probability that a susceptible person gets infected in time  $t$  and  $R(t)$  indicates the probability that an infected persons recovers in time  $t$  and becomes susceptible. These probabilities correspond exactly with the transition probabilities in the Markov model. The Markov model is calculated with a constant transition matrix since only the equilibrium should be fitted. To prevent random noise, the quantities are calculated as the mean of 20 time steps in the agent based model. This method yields the parameters  $I = 9.11\%$  and  $R = 10\%$ .

The results of the models are shown as carrier rates of the whole population. Figure 5.3 shows that these parameters indeed yield the same equilibrium in both models. The adaption phases are different due to structural differences but, as stated before, they are no matter of interest.



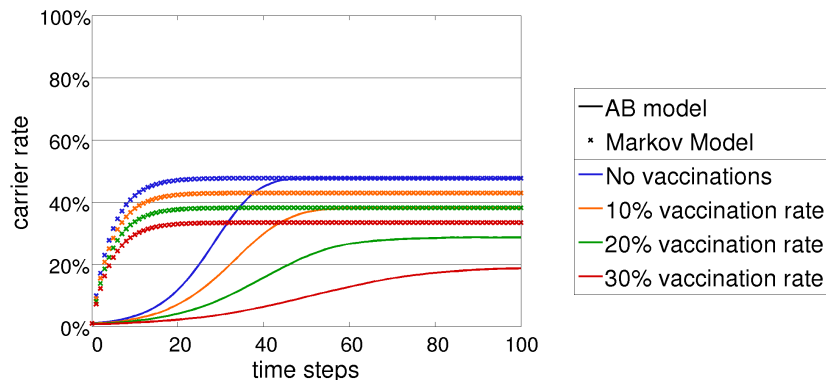
**Figure 5.3:** Carrier rates of the equally parameterized SIS models without vaccinations

### 5.2.3 Implementation of vaccinations

In both models, three different vaccinations rates are performed at the beginning of the simulation: 10%, 20% and 30%. The Markov model is parameterized based on the agent based model without vaccination. The corresponding results of these vaccinations are plotted in figure 5.4.

The Markov model shows the expected linear behavior: Vaccination of 10%, 20% or 30% of the population yields to a reduction of infected people by the same percentage. This leads to carrier rates that are significantly higher than the ones in the agent based model. The adaption phases are also different but they are not discussed here, since only the equilibrium should be compared.





**Figure 5.4:** Results of vaccinations in the Markov model and in the agent based model for SIS epidemics

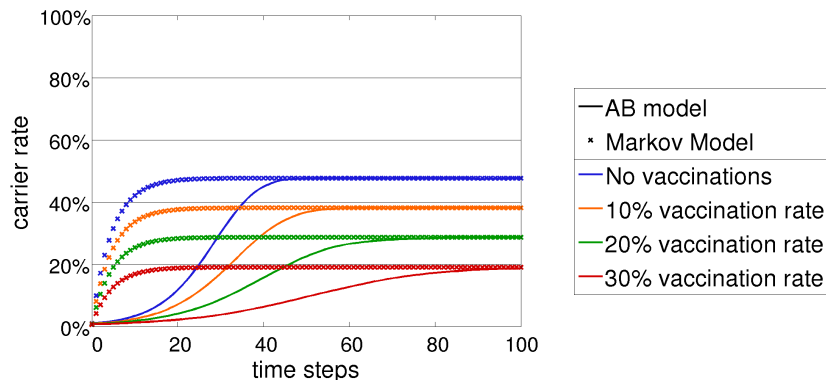
#### 5.2.4 Explanation of the differences and harmonization of the results

The background for the different results is clear because it is already meticulously shown that an epidemic, which has a propagation based on contacts, does not behave linear because the unvaccinated population is affected as well. At this point the question arises, if the Markov model can be corrected by the factor of non-linearity so that it yields carrier rates that correspond to the agent based model. This means, we are looking for a quantity that indicates how much the carrier rates are overestimated by the Markov model. Since only unvaccinated people can get infected this quantity equals exactly the benefit of the unvaccinated population, and this is the herd immunity. Hence, the herd immunity has to be calculated in the agent based model from the current carrier rate. Table 5.1 shows the results of this comparison. The herd immunity is shown as a percentage for better comparability with the other results. The small differences can be explained by random noise in the agent based model. Figure 5.5 visualizes that the carrier rates in the equilibrium in the Markov model indeed corresponds to the agent based model if they are corrected by the herd immunity.

**Table 5.1:** Comparison of carrier rates and herd immunity in the SIS model

Vaccination rate	10%	20%	30%
Carrier rate in the agent based model	38.06%	28.61%	18.96%
Carrier rate in the Markov model	42.91%	38.15%	33.38%
Divergence of the Markov model from the agent based model	11.30%	25.00%	43.20%
Herd immunity in the agent based model	11.15%	24.83%	43.00%

**This example leads to some important insights:**



**Figure 5.5:** Results of vaccinations in the Markov model (corrected by the herd immunity) and in the agent based model for SIS epidemics

- A Markov SIS model with constant transition probabilities can always be parameterized so that its equilibrium corresponds to the agent based model.
- The carrier rate in the equilibrium of a correct parameterized Markov SIS model does not correspond to the agent based model. But, decreasing the resulting carrier rate by the herd immunity yields the correct result.
- The herd immunity of a concrete SIS epidemic depends on the vaccination rate and cannot be generally set or used.
- The adaption phase - the duration, how long it takes after an intervention to reach the new equilibrium - cannot be calculated correctly with a Markov model with constant transition probabilities.

## 5.3 Comparison of the temporal progress of an SIR epidemic

### 5.3.1 Construction of the models

**The agent based model:** The agent based model is already defined in chapter 4.5.2 and is extensively examined in chapter 4.8.3. For that reason, it will not be described here again.

**The Markov model:** Like for SIS epidemics, the population of the Markov model is split into unvaccinated and vaccinated people.

**Unvaccinated people:** The Markov model for SIR epidemics is generally an extension of the SIS model. It requires three states:  $s_1 = \text{“Susceptible”}$ ,  $s_2 = \text{“Infected”}$  and  $s_3 = \text{“Resistant”}$ . The state transition diagram in figure 5.6 visualizes this model.



**Figure 5.6:** The state transition diagram of the SIR Markov model for unvaccinated people

Mathematically, the distribution of the system in a time point is represented by a 3-dimensional vector which indicates the shares of susceptible, infected and resistant people:  $\underline{a}_t = (a_{1,t}, a_{2,t}, a_{3,t})$ . The transition Matrix  $M$  shows the probabilities, that susceptible respectively infected people transfer to the state “*Infected*” (with probability  $I$ ) respectively “*Resistant*” (with probability  $R$ ). Thus, they remain in their state with probability  $(1 - I)$  respectively  $(1 - R)$ . Resistant people always remain in their state.

$$M = \begin{pmatrix} 1 - I & I & 0 \\ 0 & 1 - R & R \\ 0 & 0 & 1 \end{pmatrix}$$

**Vaccinated people:** All vaccinated people remain always vaccinated because they cannot get infected. Hence, the Markov model consists of this state “*Vaccinated*” only.

Figure 5.7 shows the tree of the whole model.

### 5.3.2 Parameterization of the model

**Parameterization of the agent based model:** The parameterization of the agent based model is the same as of the SIS model in the previous chapter 5.2.2.

**Parameterization of the Markov model:** The task is, to set the infection probability  $I$  and recovery probability  $R$  so, that the Markov model yields the same epidemic progression as the agent based model. Thus, the transition probabilities have to be set time-dependent. This is possible if the probabilities  $I$  and  $R$  are calculated as stated already for the SIS model (equation 5.1).

The infection probability  $I$  is a time-dependent function that is similar to the progression of the SIR disease (figure 5.8, left). The recovery probability is decreasing in the beginning from 10%, because due to the given recovery time many people get infected but do not recover yet. Then it increases again above 10% and finally, it ends in an almost chaotic

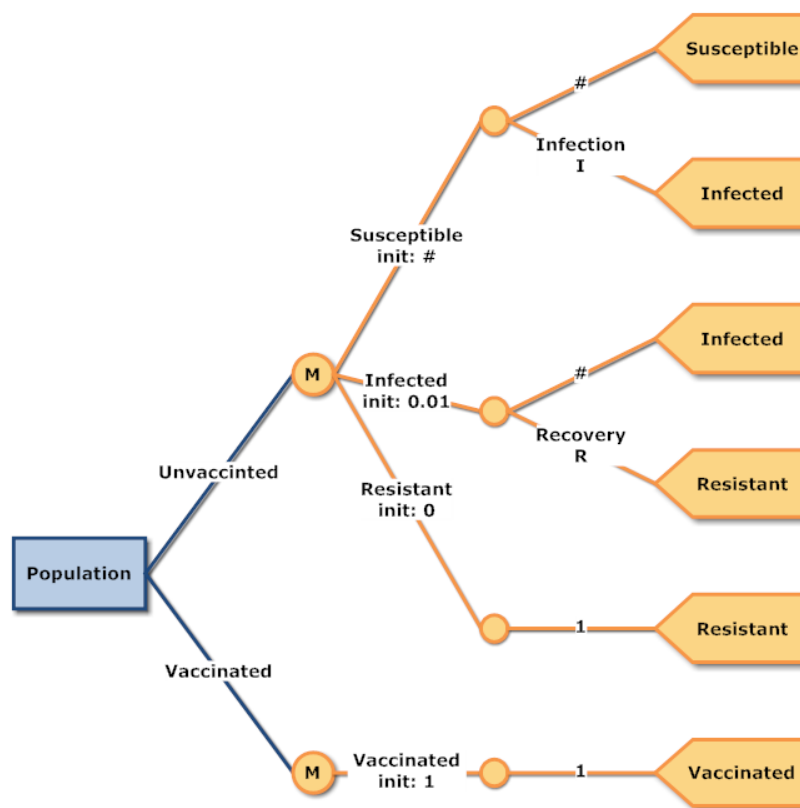
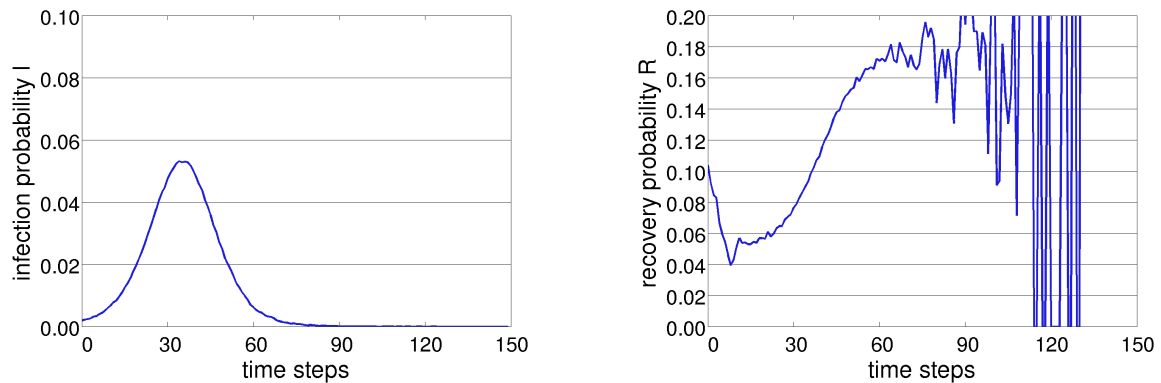


Figure 5.7: The decision tree of the SIR Markov model

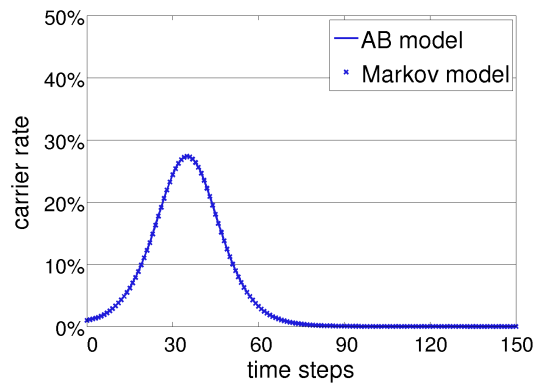
### 5.3. COMPARISON OF THE TEMPORAL PROGRESS OF AN SIR EPIDEMIC

behavior because only a few infected people are left, so that single recovering people represent high percentages (figure 5.8, right).

Using these values  $I$  and  $R$  in the Markov model leads to the same epidemic curve as the agent based model (figure 5.9).



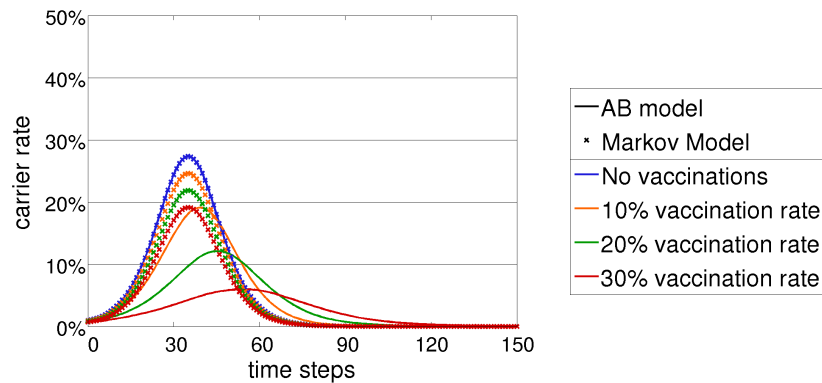
**Figure 5.8:** Time-dependent parameter in the SIR Markov model without vaccination. **Left:** Infection probability  $I$ , **Right:** Recovery probability  $R$



**Figure 5.9:** Results of the equally parameterized SIR models

#### 5.3.3 Implementation of vaccinations

A vaccination in the agent based model decreases the epidemic curve but it also distorts it, hence it takes longer until the disease becomes extinct. In the Markov model, the vaccination decreases the epidemic curve by the vaccination rate but does not change its qualitative behavior over time. Figure 5.10 shows these differences of the simulation results.



**Figure 5.10:** Results of vaccinations in the Markov model and in the agent based model for SIR epidemics

### 5.3.4 Explanation of the differences and harmonization of the results

The last step of the procedure tries to capture the error and to harmonize the different results. Vaccinated people do not participate in the epidemic curve, hence the effect is caused by unvaccinated people. This means, an effect influences the unvaccinated population that finally results in changed epidemic curves. This effect is exactly the benefit compared to the base scenario without vaccination - the herd immunity. Since the result is the temporal progression of the whole epidemic curve, the herd immunity must be calculated upon the current carrier rate (see figure 5.11).

The herd immunity represents the benefit for the unvaccinated population. To use this effect correctly the carrier rate in the Markov model needs to be multiplied by  $(1 - \text{herd immunity})$ . This means that for each time point  $t$  applies:

$$I_{\text{Markov with herd immunity}}(t) = I_{\text{Markov}}(t) * (1 - H(t))$$

with

$I(t)$  ... Carrier rate in time  $t$

$H(t)$  ... Herd immunity in time  $t$

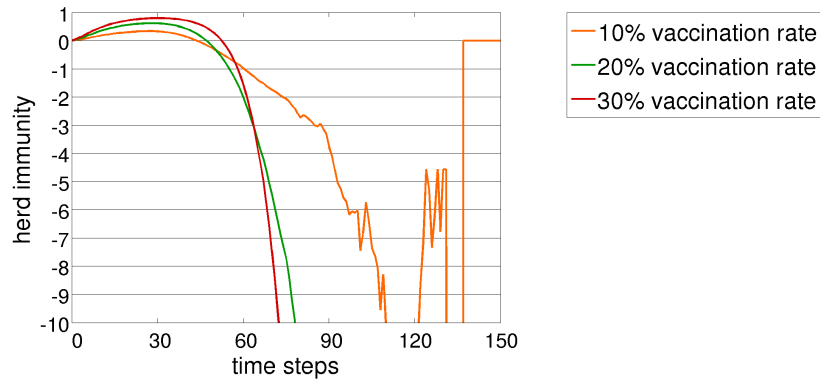
Such a recalculated carrier rate finally corresponds to the agent based model as shown in figure 5.12.

#### This example also leads to some important insights

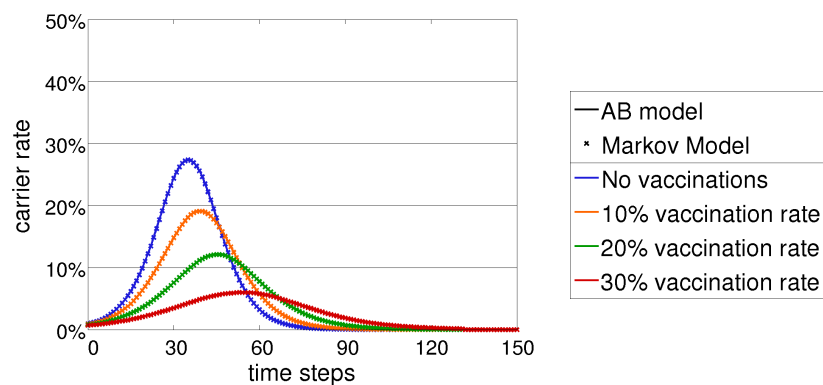
- A Markov model can always be parameterized so that its temporal progression corresponds to the agent based model.

### 5.3. COMPARISON OF THE TEMPORAL PROGRESS OF AN SIR EPIDEMIC

---



**Figure 5.11:** Herd immunity in the agent based SIR model based on the current carrier rate



**Figure 5.12:** Results of vaccinations in the Markov model (corrected by the herd immunity) and in the agent based model for SIR epidemics

- Vaccination in a correctly parameterized Markov model yields generally divergent results. Decrease of the carrier rate by the herd immunity yields the carrier rates of the agent based model.
- Herd immunity in an SIR model is not a constant parameter but a time-dependent function. Additionally, it depends on the vaccination rate and so it cannot be declared generally and has to be calculated time-dependently for every specific setting.

## 5.4 Comparison of a complete SIR epidemic

### 5.4.1 Construction of the models

This chapter requires a model that only yields the number of infections of a whole SIR epidemic.

**Agent based model:** The SIR agent based model is the same as in the previous chapter, but the result is only the total number of infections.

**Markov model:** The Markov model is a very simple structure. One time step covers one complete SIR epidemic. This allows to calculate recurring epidemics under the assumptions that no infected people are left at the end of a season. This means that for a time step only one parameter has to be set: the number of infections in the susceptible population.

### 5.4.2 Parameterization of the model

**Parameterization of the agent based model:** The parameterization is the same as in the agent based model for SIS and SIR epidemics which is set in chapter 5.2.2. A total of 78.77% of the whole population get infected during the whole epidemic in the scenario without vaccination.

**Parameterization of the Markov model:** The Markov model is parameterized by the 78.77% infected people of the agent based model. It does not contain any other parameters.

### 5.4.3 Implementation of vaccinations

The results of the agent based epidemics with vaccinations are taken from the previous chapter 5.3.3. The Markov model decreases the infections in the population by the vaccination rate due to its linearity. The results are shown in table 5.2.



**Table 5.2:** Infected population in the Markov model compared to the agent based model

Vaccination rate	Infected part of the population	
	Markov model	AB model
none	78.77%	78.77%
10%	$78.77\% \cdot 0.9 = 70.89\%$	65.09%
20%	$78.77\% \cdot 0.8 = 63.02\%$	50.85%
30%	$78.77\% \cdot 0.7 = 55.14\%$	35.32%

**General explanation of the behavior of the Markov model:** Be  $Z_{AB,p}$  the infection number of the agent based model and  $Z_{Markov,p}$  the one of the Markov model. Then the following equation holds for the Markov model:  $Z_{Markov,p} = (1-p) \cdot Z_{AB,0}$ . This means that the Markov model assumes the same infection probabilities for the unvaccinated part of the population  $(1-p)$  for any vaccination strategy.

#### 5.4.4 Explanation of the differences and harmonization of the results

Again the differences result from a benefit or disadvantage of the unvaccinated population. In this case herd immunity is calculated upon new infections since simulation start because the results are total infections of the whole epidemic. Table 5.3 shows the calculation of the herd immunity. Multiplication of the infection numbers of the Markov model by  $(1 - \text{herd immunity})$  again yields the results of the agent based model as shown in table 5.4.

**Table 5.3:** Calculation of the herd immunity in the agent based model. All calculations are performed with the exact (not rounded) values.

Vaccination rate	Herd immunity
10%	$\frac{78.77 - 65.09}{78.77} = 0.0819$
20%	$\frac{78.77 - 50.85}{78.77} = 0.1930$
30%	$\frac{78.77 - 35.32}{78.77} = 0.3595$

**Table 5.4:** Results of the Markov model with herd immunity compared to the agent based model

Vaccination rate	Infected part of the population	
	Markov model with herd immunity	AB Model
none	78.77%	78.77%
10%	$70.89\% \cdot (1 - 0.0819) = 65.09\%$	65.09%
20%	$63.02\% \cdot (1 - 0.1930) = 50.85\%$	50.85%
30%	$55.14\% \cdot (1 - 0.3595) = 35.32\%$	35.32%

## 5.5 Analytic considerations

The relationship between model results and herd immunity, which is already motivated by the three examples, can be analytically shown as well.

Let  $E_{model,p}$  be the result of an epidemic model with vaccination rate  $p$  in a point in time or over a period of time. If the Markov model and the agent based model are equally parameterized in the base scenario without vaccination to yield the same results, then one has:  $E_{Markov,0} = E_{AB,0}$ .

The Markov model in case of vaccination assumes that vaccinated people cannot have any disadvantage, and unvaccinated persons have the same result as before in the base scenario. This means:  $E_{Markov,p} = (1 - p) \cdot E_{Markov,0}$ .

Let  $H_p$  be the herd immunity to a vaccination rate  $p$  and it should be calculated upon the same disadvantage and time as the result of the model  $E_{Markov,p}$ . The herd immunity is calculated in the agent based model. Then we can state by definition that:

$$H_p = \frac{E_{AB,0} - \frac{E_{AB,p}}{(1-p)}}{E_{AB,0}}$$

This leads to:

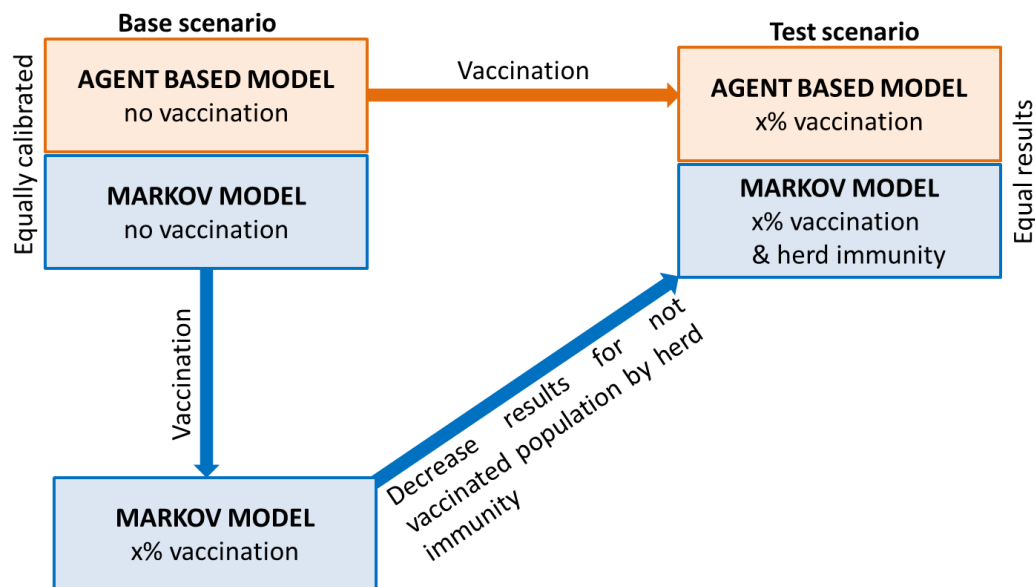
$$\begin{aligned} E_{Markov,p} \cdot (1 - H_p) &= E_{AB,0} \cdot (1 - p) \cdot \left( 1 - \frac{E_{AB,0} - \frac{E_{AB,p}}{(1-p)}}{E_{AB,0}} \right) \\ &= E_{AB,0} \cdot (1 - p) \cdot \frac{\frac{E_{AB,p}}{1-p}}{E_{AB,0}} \\ &= E_{AB,p} \end{aligned}$$

**Consequence:** This consideration shows that the herd immunity, if calculated upon an appropriate disadvantage and time, always completely describes the differences between Markov models and contact-based models.

## 5.6 Conclusions

The results of this chapter clearly show the the highly different behavior of linear, stochastic models and contact based models which are represented by Markov models and agent based models. Examples show both qualitative and quantitative differences between these approaches. Application of the already introduced herd immunity make these differences graspable and they can be explained and calculated in a standardized way. This leads to a few very important insights. These relationships are also shown in figure 5.13.

- A Markov model with time-dependent transition probabilities can always be parameterized or calibrated that its results fit contact based models or other data. For comparability, it is important that both models produce results of the same structure.
- Correctly parameterized Markov models yield qualitative and quantitative different results compared to contact based models upon interventions on the system (for example vaccinations).
- Appropriate calculation of herd immunity in the contact based model completely describes the differences of the results of the two model types. The disadvantage in the Markov model, decreased by the herd immunity, yields the same results as the contacts based model.
- Herd immunity represents the impact of contacts on the propagation of an epidemic compared to fixed, independent infection probabilities.
- The right calculation of herd immunity depends on the used Markov model.
- Herd immunity cannot be generally declared for epidemics. It is highly dependent on the whole configuration of the system and on the intervention.
- Generally, herd immunity is not constant but a time-dependent function.



**Figure 5.13:** Relationship between agent based models and Markov models through herd immunity

As a consequence, this chapter shows that consideration of contacts is a necessary requirement for modeling of infectious diseases.

CHAPTER 5. COMPARISON OF MARKOV MODELS AND AGENT BASED  
MODELS FOR EPIDEMIC SIMULATIONS

---

# Chapter 6

## Simulation of two competitive serotypes

### 6.1 Consideration of serotypes

In the previous chapters, we always considered a single kind of a pathogen. However, pathogens are developing and changing. Such changes occur because of different reasons like mutations or genetic recombinations. Thus, new types of a pathogen emerge which are called “strains” or “serotypes” [35].

Due to varying characteristics of different serotypes, it is often necessary to differentiate and consider them separately.

- Different serotypes can cause different reactions of the human body. For example some serotypes might cause severe progresses of the disease while others only cause a mild disease.
- Vaccines often do not work against all kinds of a pathogen but only against a few serotypes. This fact can lead to serious consequences, as the following chapters show. Conjugate vaccines serve as an example. They work by recognizing bacteria through their polysaccharide covering. Different strains of a pathogen often do not have the same covering and so some of them cannot be recognized and the vaccine does not work against them. Well known examples of conjugate vaccines are those against streptococcus pneumoniae or haemophilus influenzae.
- Serotypes of a pathogen can affect each other so that they cannot spread freely any more. This means that they have to share their living space which are potentially susceptible persons.

## 6.2 Competition of serotypes

Existing or newly emerged serotypes often do not exist independently, but instead they affect each other. Hence, this leads to a competitive behavior which results in displacement or replacement effects.

Such a competitive behavior, or at least its results, can be observed in several studies as the following examples show.

### 6.2.1 *Streptococcus pneumoniae*

*Streptococcus pneumoniae* are bacteria that can cause pneumonia, otitis media (middle ear inflammation), meningitis and sepsis (blood poisoning) especially among children and old people. There are over 90 serotypes which can be distinguished by its outer covering. The prevalences of these serotypes differ a lot. Existing vaccines are conjugate vaccines which work only against certain serotypes [9].

**Direct competition:** A US-American study from 2010 examined simultaneous occurrence of the two pneumococcal serotypes 24F and 7 in mice, under special consideration of simultaneous occurrence of *haemophilus influenzae* bacteria. The results show clearly that the two pneumococcal serotypes cannot exist independently from each other within a mouse because in different situations serotype 24F or serotype 7 is ousted. Hence, the competition is obvious. [36]

**Effects of the competition:** A common vaccine against *streptococcus pneumoniae* is Prevenar, also called PCV7, which is in the market since 2004. PCV7 works against 7 frequently occurring serotypes that are altogether 80% of all pneumococcal occurrences in the USA. First studies after it entered the market showed a strong decrease of pneumonia and otitis media cases. However, they often only considered sicknesses caused by one of those 7 serotypes [39]. But new studies from 2009 and 2010 reveal a strong increase of sicknesses by other serotypes while the occurrence of PVC7-serotypes was decreasing. Altogether, they found only a small reduction or even no change of the number of cases as an effect of vaccinations with PCV7 [22] [19].

Hence, pneumococcal serotypes must influence each other if a decimation of some serotypes cause an increased prevalence of other serotypes. If this influence would not exist, there would not be an increase of sicknesses caused by other serotypes as an effect of the vaccinations.

### 6.2.2 Influenza

The influenza, also called flu, is an infectious disease that is caused by viruses of the family of the Orthomyxoviridae. Influenza viruses are classified into the genera A, B and

C while A is classified into serotypes. Further variations and mutations mostly exist within serotypes of Type A. Hence, a separate consideration of single serotypes or variations of influenza viruses is inevitable. [2] [8]

In 2009, the so called “swine flu” appeared which is a variation of the H1N1-influenza virus A. This virus is described to be very infectious. Studies show that the common “seasonal influenza viruses” have been widely ousted by the swine flu virus in the influenza season 2009/2010. Mutual influence of these serotypes is a plausible explanation of this behavior. [1]

#### 6.2.3 Haemophilus Influenzae

Haemophilus Influenzae are bacteria that can cause inflammations of the respiratory tract and, especially for children, meningitis and other inflammatory diseases [58] [36]. Based on their outer covering, these bacteria are classified into six types a,b,c,d,e and f [26]. A conjugate vaccine against Haemophilus Influenzae type b was introduced in the year 2000 but, in contrast to streptococcus pneumoniae, no impact on the other serotypes has been observed [32] [34].

#### 6.2.4 Serotype shift and consequences of competitive serotypes

Studies show that different serotypes sometimes influence each other by sharing and competing for living space which are potential colonizable people.

Fighting single serotypes of a pathogen might lead to a situation where other serotypes have suddenly a larger living space available and can occupy the places of the defeated serotypes. Such an increasing occurrence of strains caused by interventions like vaccinations is called “serotype shift” or “serotype replacement”.

**Consequences:** Interventions on systems of competitive serotypes always result in a redistribution of the living space. This process is time-dependent and nontrivial; its result usually depends on several factors. The following chapters try to examine, describe and explain such effects.

### 6.3 Models with competitive serotypes in literature

A literature search reveals a few attempts and approaches to simulate competitive serotypes.

One of the first models originates from K. Dietz in the year 1979. It is an SIR model that consists of differential equations with two serotypes under the assumption that persons cannot be infected with both serotypes at one time. This study extensively examines

equilibriums and stability conditions. Vaccinations are not considered, hence phenomena like herd immunity and serotype shift are not prevalent. [11]

The ideas of Dietz have been picked up and further developed several times, hence most models of competitive serotypes are realized by differential equations or system dynamics.

S. Gupta examined in 1997 a differential equations model with four serotypes and the possibility of vaccinations with optional cross protection. The main part of the work is examination of equilibriums. Serotype shift in case of vaccinations is clearly visible as well. [20]

Also in 1997, M. Lipsitch presented a system dynamics models for competitive serotypes. People in this model can be infected with one or both serotypes and can recover. In all compartments, he uses a constant death rate. The number of new susceptible people that enter the model equals to the died people. Hence, the population size remains constant. The flow rates to the infected-compartments do not depend on the number of vaccinated people, hence herd immunity is not considered and cannot emerge. However in case of vaccinations strong serotype shifts occur. [33]

In 2002, Lipsitch thoroughly examines the term “serotype shift”. In the publication he states that a pairwise interaction between serotypes is a necessary requirement for occurrence of serotype shift. [35]

Lipsitch’s modeling approach has been further practiced by other research groups. In 2002 A. Scherer and A. McLean examined a system dynamics model of an SIRS epidemic that result in an oscillating state and is able to create effects of serotype shift. [49]

Another interesting study was published in 2003 by L. Esteva and C. Vargas about dengue fever. In general they use an SIR model with two serotypes. In addition to the human population they model a mosquito population with infected and healthy mosquitos while infection rates of humans and mosquitos depend on each other. Furthermore a constant rate of persons die. They do not leave the model completely but re-enter the model as newborn susceptible people. Because of deaths and births, the model leads to a stable system with equilibriums, even though it is an SIR model. The main part of the study deals with examination of these equilibriums and statements about stability. [13]

Yidi Zhang also picked up the approach of Lipsitch 2006 in her dissertation, extended them consequently to different SIS and SIR models with and without vaccinations and examined the new systems by analytic aspects. [64]

Summarizing this, we can say that all found studies use differential equations or system dynamics. This might be, because traditional statistic or stochastic models are not sufficient for simulation of competition of serotypes. Furthermore, such differential equation systems can be examined analytically very well. The different studies show interesting approaches, ideas and also the limits of this methodology. While serotype shift is often clearly visible, herd immunity is often deliberately avoided or ignored. However, the focus of the found studies is not examination and explanation of occurring effects but rather statements about equilibriums and stability.



## 6.4 Preliminary remarks on modeling of competitive serotypes

In chapter 3, a few approaches have been presented for modeling and simulation of the propagation of an epidemic. Pathogens or serotypes that do not influence each other can still be simulated with these models by calculating every pathogen independently in a separate model.

But in case of competitive serotypes, models are required that calculate the share of a living space. They also need to take account of the redistribution of this space after interventions on the system [35].

**Concrete modeling:** The assumptions for modeling of competitive serotypes are a priori not clear and cannot be clarified completely by medical studies. A competition process might even differ for different pathogens. Hence, the following chapters show different possibilities for modeling of such a competition.

Again, an agent based model is used as an approach for simulation of competitive serotypes and for examination of herd immunity and serotype shift. This method has been used successfully for single serotypes because it models epidemics as a result of interactions of persons and thereby causing transmissions. Furthermore, it is flexibly extensible and easy to handle (for more information on this topic read also the motivation of usage in chapter 4.5).

**Assumptions on agent based models:** The assumptions on agent based models that will be used for simulation of competitive serotypes build on the ones for single serotypes, which are described in chapter 4.6.3.

- The agents in the model are persons.
- All persons in the model are of the same type. This means that they have the same behavior and attributes and only differ in their temporal state.
- Infection states and progression of an infection: defined separately for each case.
- Contacts are uniformly distributed among the whole population. The parameter  $k$  = “contacts per person per time step” defines the number of contacts.
- The model is calculated in discrete time steps. Updates of infection states of persons happen simultaneously for the whole population once per time step.

**Strength of serotypes:** Epidemic parameters can be set for every serotype separately. For simple SIS and SIR models, these parameters are the infection probability and the recovery time. It is possible to distinguish two fundamental cases:

- **Equal parameters:** If both serotypes have the same settings for all epidemic parameters, then one can expect the same qualitative and quantitative behavior.
- **Unequal parameters:** If the value of at least one epidemic parameter differs for both serotypes, then one can generally expect a different qualitative or quantitative behavior of the serotypes. How these differences occur cannot be stated in general and has to be examined in concrete situations.

## 6.5 Modeling approach 1: Competition by carrying only one serotype

This approach for modeling of competitive serotypes is based on the following three assumptions:

- A person can be infected with a maximum of one serotype in a time point.
- A person that is infected with a serotype cannot be infected with another serotype.
- The first two assumptions do not describe the system completely. The model simulates discrete time steps where contacts happen and people get infected simultaneously at the end of each time step. Contacts are not considered to be in order.  
**Problem:** What happens if a susceptible person would get infected with both serotypes upon two different contacts?  
**Solution:** A person remembers all infections that happen within a time step without changing their infection state. At the end of the time step, when the states are updated, the person chooses randomly one of the infection.  
**Consequence:** If a person gets infected  $x$  times with serotype 1 and  $y$  times with serotype 2 within a single time step, then the probabilities that they become infected with serotype 1 or 2 in the end correspond to the ratio  $x : y$ .

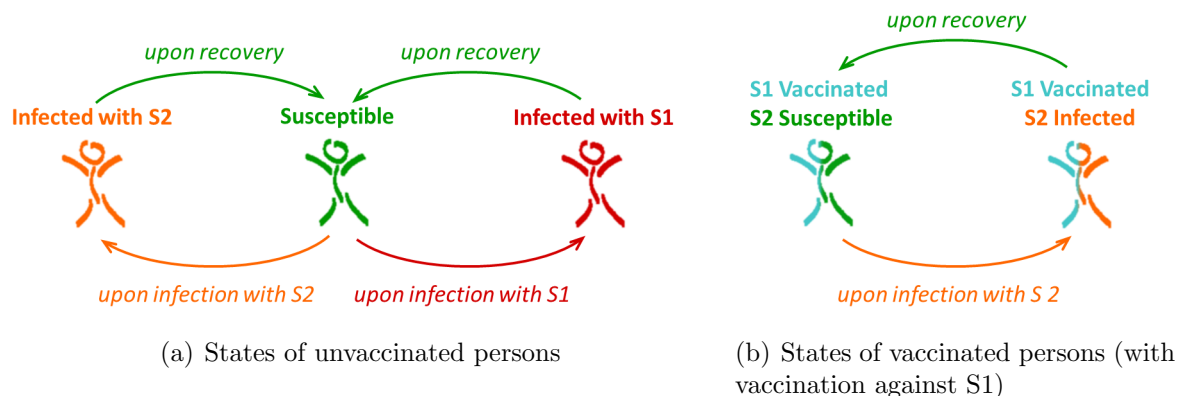
**Simulations:** The competitive behavior of two serotypes is simulated for both SIS and SIR models with different parameters to study this approach. Epidemics with competitive serotypes produce a stronger random noise, hence 5 Million people will be simulated.

- Number of persons:  $N = 5$  million
- Contacts per person per time step:  $k = 10$
- Start values: In the initial state 1% of the population is infected with serotype 1 and serotype 2 each and 98% are susceptible.
- Three scenarios with different epidemic parameters are simulated:
  1. **Same epidemic parameters:**

- Infection probability:  $p = 0.02$
  - Recovery time:  $d = 10$  time steps
2. **Different infection probability:**
- Infection probability:  
 $p_1 = 0.022$  for serotype 1 and  $p_2 = 0.02$  for serotype 2
  - Recovery time:  $d = 10$  time steps for both serotypes
3. **Different recovery time:**
- Infection probability:  $p = 0.02$  for both serotypes
  - Recovery time:  $d_1 = 11$  for serotype 1 and  $d_2 = 10$  for serotype 2

### 6.5.1 Modeling an SIS epidemic with approach 1

**Description:** In an agent based SIS model, susceptible people can be infected by infected people and become susceptible again after recovery. Figure 6.1 shows the process of unvaccinated persons as well as of vaccinated persons.

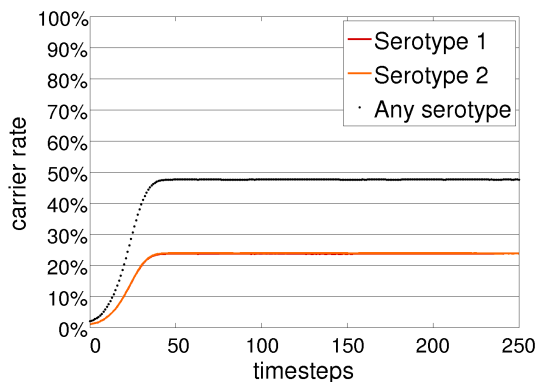


**Figure 6.1:** Visualization of an SIS disease with two competitive serotypes (using approach 1)

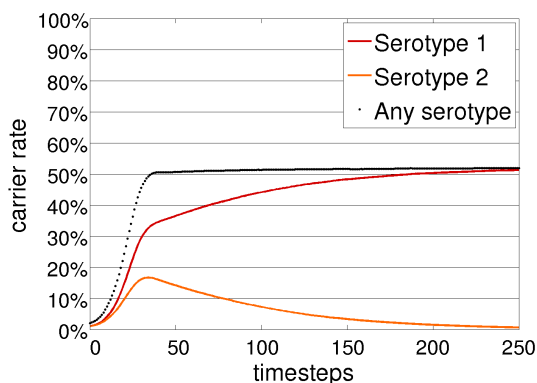
**Simulation results:** It turns out that in case of equivalent epidemic parameters, the model reach an equilibrium where both serotypes exist on the same level. Figure 6.2 shows the results of the simulation.

Different parameters for the serotypes lead to distinction of the weaker serotype. The surviving serotype then exists like in the single-pathogen-model and reaches an equilibrium. Interestingly, the number of carriers is constant all the time even during the extinction phase of the weak serotype. Figures 6.3 and 6.4 show the results of the simulations.

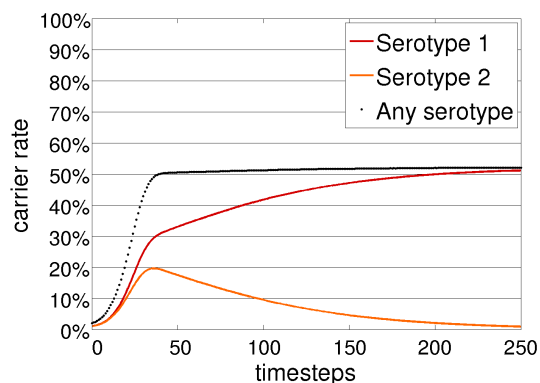
**Comparison with differential equations:** Such a system can also be modeled and simulated with differential equations as shown in formula 6.1.



**Figure 6.2:** Carrier rates of an SIS epidemic with equivalent serotypes and approach 1 ( $p = 0.02$ ,  $d = 10$  time steps)



**Figure 6.3:** Carrier rates of an SIS epidemic with different serotypes and approach 1 ( $p_1 = 0.022$ ,  $p_2 = 0.02$ ,  $d = 10$  time steps)



**Figure 6.4:** Carrier rates of an SIS epidemic with different serotypes and approach 1 ( $p = 0.02$ ,  $d_1 = 11$  time steps,  $d_2 = 10$  time steps)

$$\begin{aligned}
 \dot{S}(t) &= -\alpha_1 \cdot I(t) \cdot S(t) - \alpha_2 \cdot J(t) \cdot S(t) + \beta_1 \cdot I(t) + \beta_2 \cdot J(t) \\
 \dot{I}(t) &= \alpha_1 \cdot I(t) \cdot S(t) - \beta_1 \cdot I(t) \\
 \dot{J}(t) &= \alpha_2 \cdot J(t) \cdot S(t) - \beta_2 \cdot J(t)
 \end{aligned} \tag{6.1}$$

with the initial values  
 $I(0) = I_0 \geq 0, J(0) = J_0 \geq 0, S(0) = N - I(0) - J(0)$

With:  $S(t)$  ..... Number of susceptible people in time t  
 $I(t), J(t)$  ..... Number of infected people with serotype 1  
 respectively 2 in time t  
 $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$  .... Model parameter  
 $N$  ..... Number of persons

Christoph Urach examined such a system of differential equation in his master's thesis. He revealed analytically that - like in the agent based model - no equilibriums for unequal serotypes can be found, except for the trivial ones when one serotype becomes extinct. Moreover, he found out that two equivalent serotypes lead to an equilibrium that is an unstable solution of the equations. [54]

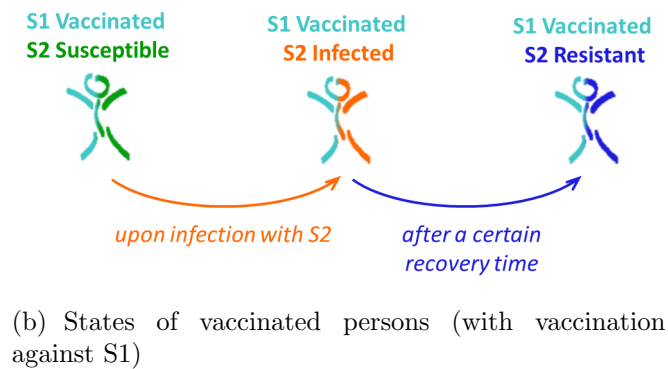
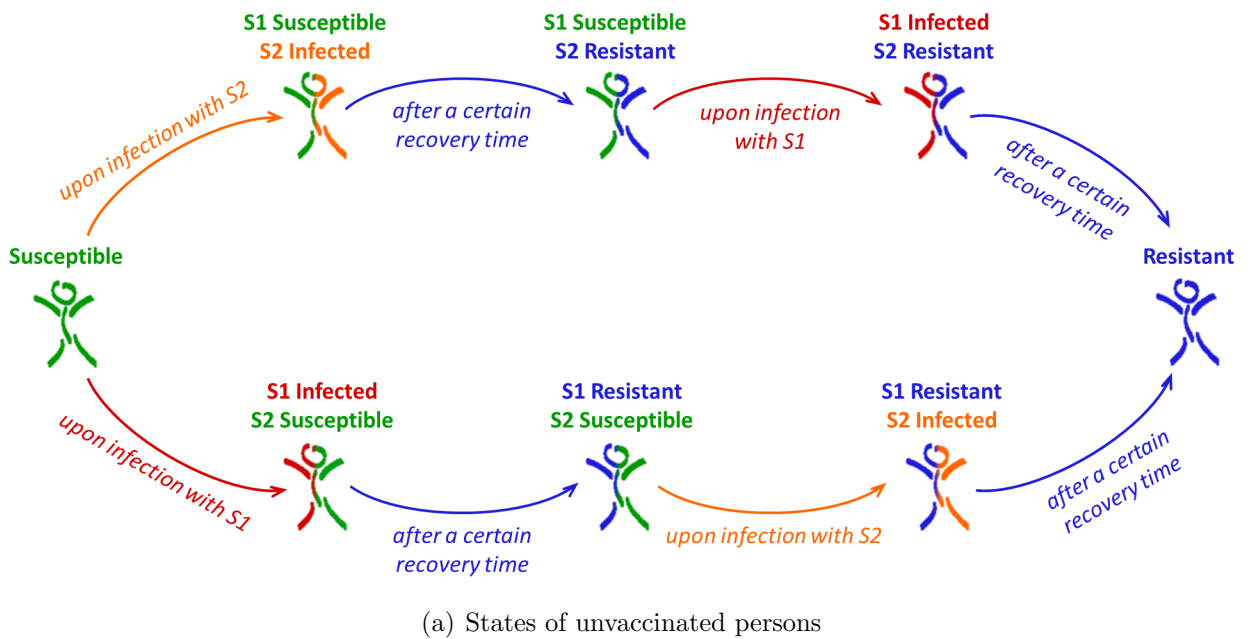
**Conclusions:** This method is not appropriate for simulation of competitive serotypes in SIS epidemics because it is not possible to reach an equilibrium or at least coexistence of unequal serotypes since one serotype always becomes extinct. Equal serotypes can survive in the same equilibriums but this system is unstable. Then, even minimal influences or changes lead to extinction of one serotype.

### 6.5.2 Modeling an SIR epidemic with approach 1

**Description:** An SIR model with competitive serotypes turns out to be much more complex than an SIS model. Susceptible persons can potentially get infected with any serotype. But it is not enough to indicate if a person is infected with a serotype, because it is relevant if a person is already resistant against the other serotype due to a prior infection. This is because a person get get infected again only after recovery of the first infection. Then, after recovering from the second infection, a person is completely resistant. Figure 6.5 shows this situation for vaccinated and unvaccinated persons.

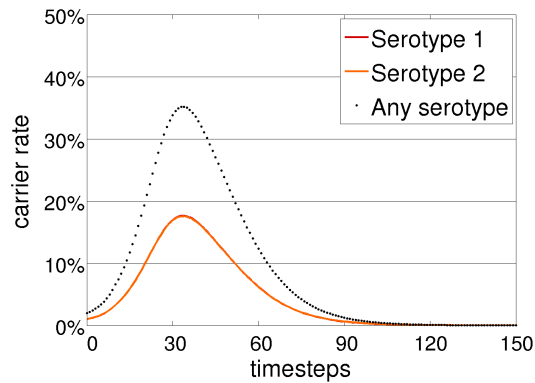
**Simulation results:** Equal serotypes lead to the well known epidemic curves in line in simple SIR models for both serotypes. Figure 6.6 shows the results.

If we change the infection probability or the recovery time of one serotype, then only the stronger serotype spreads like a normal SIR epidemic. The weaker serotype has a lower peak but has an epidemic curve that is distorted to the right and lasts longer. The

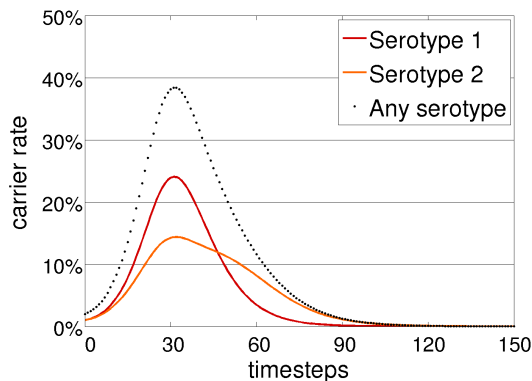


**Figure 6.5:** Visualization of an SIR disease with competitive serotypes (using approach 1)

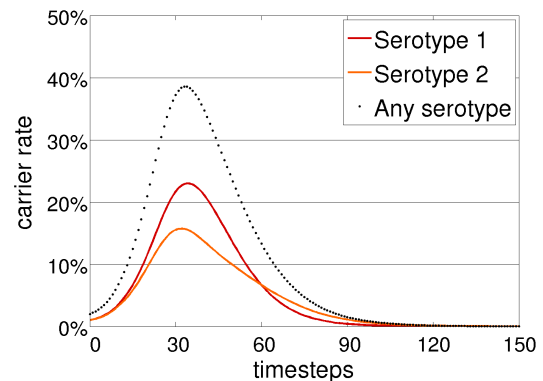
spread of the weaker serotype seems to be highly dependent on the actual parameters. The epidemic curves are shown in figures 6.7 and 6.8.



**Figure 6.6:** Carrier rates of an SIR epidemic with equivalent serotypes and approach 1 ( $p = 0.02$ ,  $d = 10$  time steps)



**Figure 6.7:** Carrier rates of an SIR epidemic with different serotypes and approach 1 ( $p_1 = 0.022$ ,  $p_2 = 0.02$ ,  $d = 10$  time steps)



**Figure 6.8:** Carrier rates of an SIR epidemic with different serotypes and approach 1 ( $p = 0.02$ ,  $d_1 = 11$  time steps,  $d_2 = 10$  time steps)

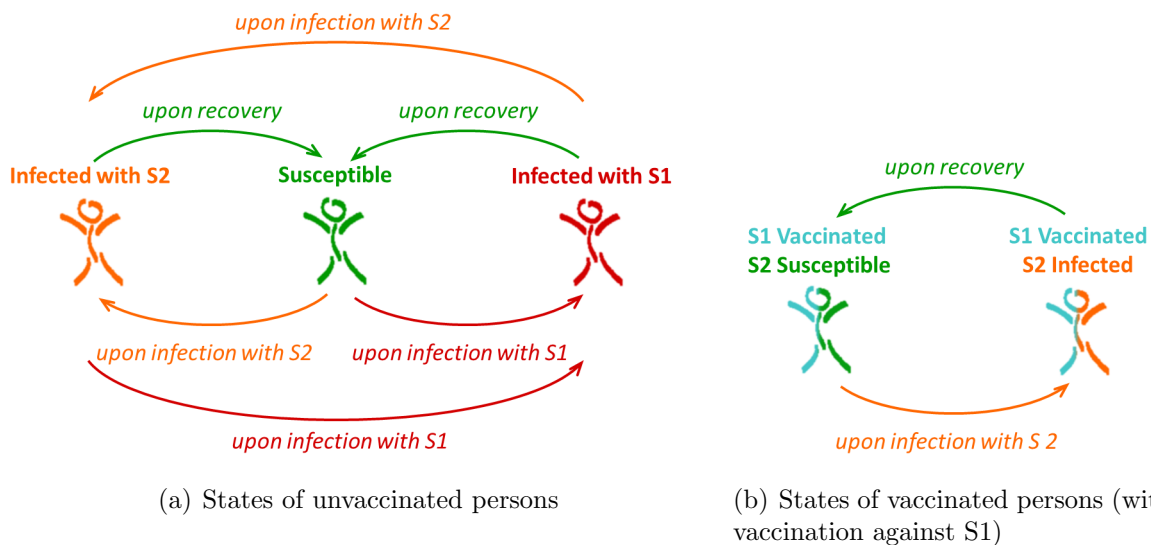
**Conclusions:** The model leads to well known SIR epidemic curves for equal parameterized serotypes and a distorted curve of one serotype in case of unequal parameterization. However, the models seems to yield usable results.

## 6.6 Modeling approach 2: Competition by carrying only one serotype and mutual susceptibility

The first modeling approach shows a behavior that might not meet the requirements. Especially for SIS epidemics, a stable coexistence of both serotypes is commonly required and for SIR epidemics an undistorted curve for both serotypes might be desired.

The idea in the second approach is that infected persons can still be infected with the other serotype. Since it is still required that people can carry at most one serotype, they have to lose the initial serotype as soon as they get infected with the other one. The background of this idea is a study that already revealed such processes for streptococcus pneumoniae [36].

Figures 6.9 and 6.10 show visualizations of this approach. The following SIS and SIR epidemics are simulated with the same parameters as in chapter 6.5.



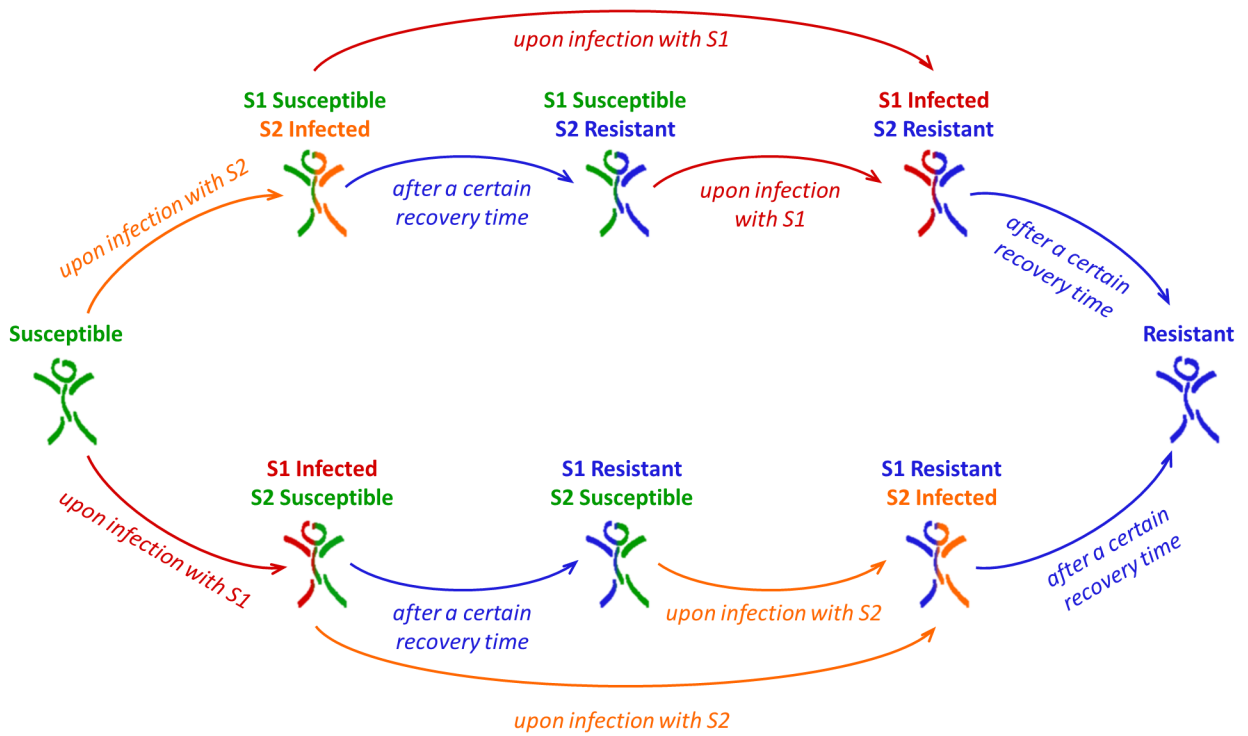
**Figure 6.9:** Visualization of an SIS disease with two competitive serotypes (using approach 2)

### 6.6.1 Modeling an SIS epidemic with approach 2

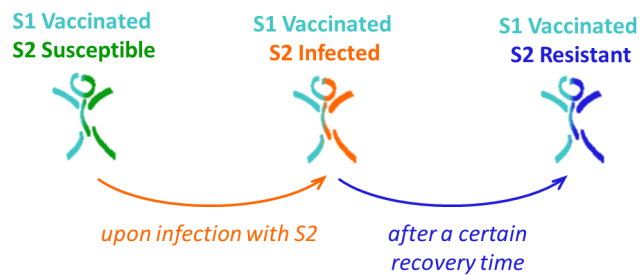
This approach finally allows stable coexistence of competitive serotypes. The system reaches a stable equilibrium after a short adaption phase, independently of the start values. Calibration of various levels seems to be possible with appropriate choice of the infection probability and the recovery time. Figures 6.11, 6.12 and 6.13 show the carrier rates of the test simulations.



6.6. MODELING APPROACH 2: COMPETITION BY CARRYING ONLY ONE SEROTYPE AND MUTUAL SUSCEPTIBILITY

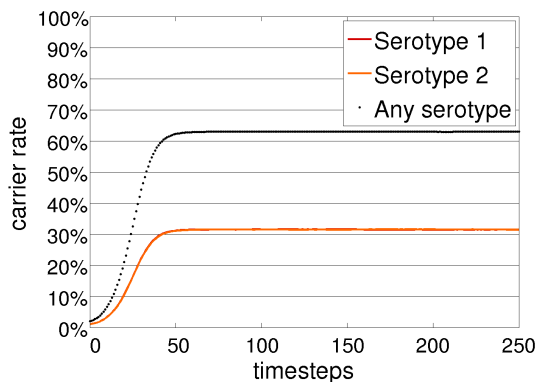


(a) States of unvaccinated persons

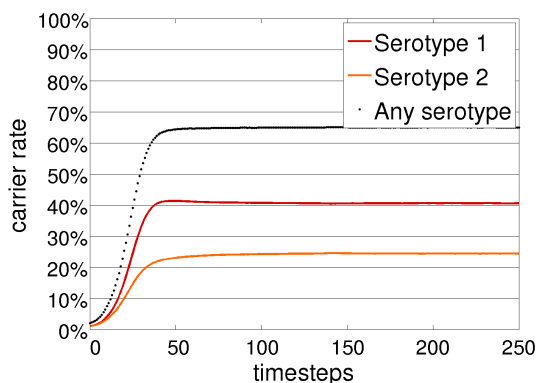


(b) States of vaccinated persons (with vaccination against S1)

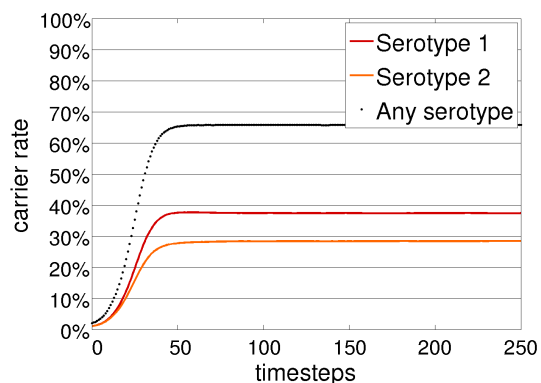
**Figure 6.10:** Visualization of an SIR disease with two competitive serotypes (using approach 2)



**Figure 6.11:** Carrier rates of an SIS epidemic with equivalent serotypes and approach 2 ( $p = 0.02$ ,  $d = 10$  time steps)



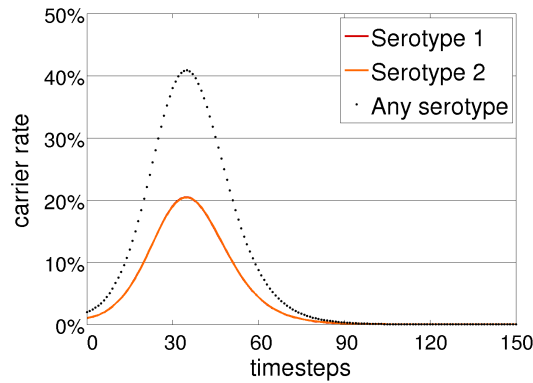
**Figure 6.12:** Carrier rates of an SIS epidemic with different serotypes and approach 2 ( $p_1 = 0.022$ ,  $p_2 = 0.02$ ,  $d = 10$  time steps)



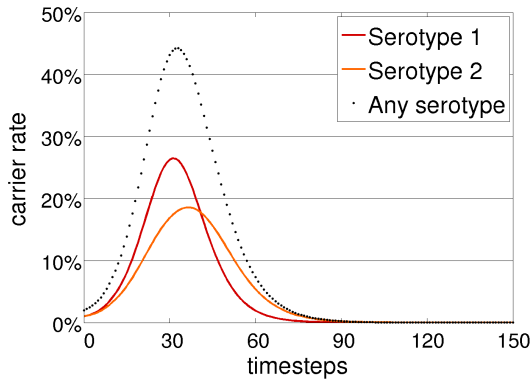
**Figure 6.13:** Carrier rates of an SIS epidemic with different serotypes and approach 2 ( $p = 0.02$ ,  $d_1 = 11$  time steps,  $d_2 = 10$  time steps)

### 6.6.2 Modeling an SIR epidemic with approach 2

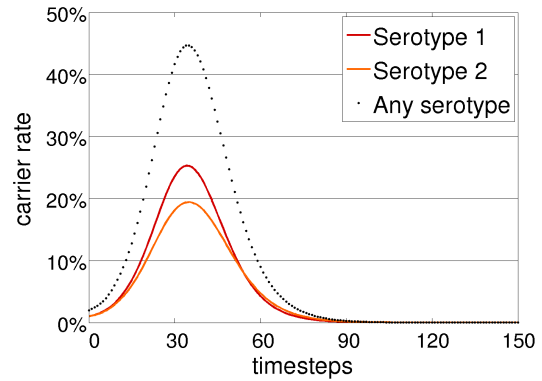
This approach yields epidemic curves that look like normal SIR epidemics for both serotypes, even if they are not equally parameterized. Depending on the settings of infection probability and recovery time, the epidemics of each of the serotypes last different periods of time and have different peaks. Figures 6.14, 6.15 and 6.16 show the results of the test simulations.



**Figure 6.14:** Carrier rates of an SIR epidemic with equivalent serotypes and approach 2 ( $p = 0.02$ ,  $d = 10$  time steps)



**Figure 6.15:** Carrier rates of an SIR epidemic with different serotypes and approach 2 ( $p_1 = 0.022$ ,  $p_2 = 0.02$ ,  $d = 10$  time steps)



**Figure 6.16:** Carrier rates of an SIR epidemic with different serotypes and approach 2 ( $p = 0.02$ ,  $d_1 = 11$  time steps,  $d_2 = 10$  time steps)

### 6.6.3 Conclusions of modeling approach 2

The changes from approach 1 to 2 are able to yield the desired effects: SIS epidemics have stable, coexisting serotypes in an equilibrium and adjustable levels of both serotypes. The

results for SIR epidemics are epidemic curves of both serotypes that are not distorted any more, but both look like the known curves of simple SIR epidemics with one serotype.

This modeling approach is not necessarily more correct than the first one. Actually, the point is to show approaches and methods to cover a wide range of various epidemic propagations. The decision on which approach is better depends on the concrete disease and has to be made case-dependent.

# Chapter 7

## Herd immunity and serotype shift in case of two competitive serotypes

### 7.1 Initial situation

The model structures show that the situation is much more complex for competitive serotypes than for single pathogens. For single serotypes, only the occurrence of the pathogen is relevant. Now, one has to differentiate here between the two serotypes and keep in mind that the sum of the disadvantages by any serotype might be relevant as well.

The goal of this chapter is qualitative and quantitative representation of the impact of vaccinations and explanation of possibly occurring effects. At first, these intentions require exact definitions of the used terms. Then, they can be examined.

### 7.2 Requirements and preliminary remarks

The following assumptions are made generally for the remarks, definitions and calculations, therefore will not be mentioned explicitly every time:

- There are two serotypes that are in competition with each other. How this competition works is irrelevant for definitions and theoretic considerations of epidemiologic effects. However, competition will be stated clearly for simulation of the model.
- The vaccination only works against one serotype. The infection probability and the personal disease progression of the other serotype is not affected by the vaccination.
- Another assumption is that the vaccination completely works for all vaccinated people. This means that they get immune against the vaccinated serotype. In reality, there are persons who do not get immune by the vaccinations. They cause costs but for epidemiology they do not need to be considered separately because they remain susceptible like unvaccinated persons.

## CHAPTER 7. HERD IMMUNITY AND SEROTYPE SHIFT IN CASE OF TWO COMPETITIVE SEROTYPES

---

- The serotype against which the vaccination works will be called “*serotype V*” (“V” like “vaccinated”). Analogously, the other serotype against which the vaccination does not work will be called “*serotype N*” (with “N” for “not vaccinated”).

The definitions and calculations of this chapter follow the ones of the single-pathogen-case because they proved themselves successfully and turned out to be sufficiently flexible. It definitely makes sense to calculate effects as a comparison between a base and a test scenario. Again, it is using a general view of the disadvantage that need to be set for concrete situations, and should provide enough flexibility.

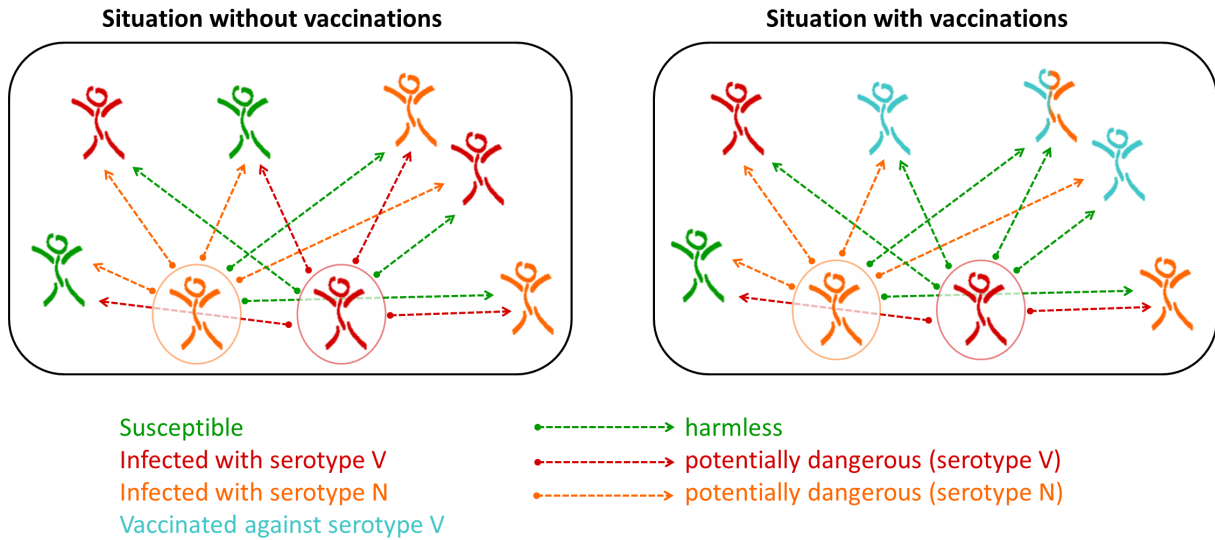
**Herd immunity:** Changes of carriage with serotype V only affects the unvaccinated population because vaccinated persons cannot be infected at all with this serotype. Hence, it is obvious to call the effect on this serotype herd immunity and calculate it analogously to the single-serotype-case.

**Serotype shift:** Dealing with Serotype N is more complex. People can be infected by these pathogens even after a vaccination, so it always affects the whole population. This leads to the following general suppositions:

- Vaccination against serotype V can potentially lead to an increasing occurrence of serotype N because it has a larger living space due to less competition.
- If there is an increasing occurrence of serotype N upon a vaccination, then it affects the whole population.
- Such an effect of increasing occurrence of serotype N can be interpreted as “serotype shift”.
- The vaccinated and the unvaccinated part of the population might be differently affected by the serotype shift because their conditions are not the same.
- Hence serotype shift might be calculated optionally for the whole population or separately for the vaccinated and unvaccinated part.

The causes of serotype shift are visualized in figure 7.1. It is important to consider that the vaccination does not hinder the propagation of serotype N by any means. But, serotype N does not need to compete with serotype V for vaccinated persons any more. The figure shows this fact: In the right picture there are two green arrows which are green in the left picture. Therefore, serotype N can spread more freely and hence it can be stronger.

**General changes:** Potentially interesting values are general changes of pathogen occurrence, independently of the serotype. For example, this could be the sum of carriers with serotype N and serotype V. The unvaccinated population part can expect a decrease



**Figure 7.1:** Visualizing the emergence of serotype shift

in disadvantages due to herd effects while they do not get in touch with the intervention. Hence, it can be seen as some kind of herd immunity. For the vaccinated population part it might mean a decrease due to immunity of serotype V. But, if vaccinated people expect an unchanged infection rate with serotype N and experience an increase, then the condition is called serotype shift. However, serotype-independent effects are not any additional effects. Instead, they only the sum of serotype N and V and therefore compose of herd immunity and serotype shift. On the whole population, this effect is based on the sum of both serotypes and it is composed of herd immunity and serotype shift. Hence, it does not need to be calculated separately. However, these effects can still be of important interest.

### 7.3 Definitions of herd immunity and serotype shift

The disadvantage is generally named  $T$ . In the case of two competitive serotypes, it needs three indexes:

$$T_{\{\text{affected part of the population}\},\{\text{base or test scenario}\},\{\text{serotype N or V}\}}$$

The affected parts of the population are indicated by  $N$  (for “not vaccinated population”),  $V$  (for “vaccinated population”) and  $A$  (for “all”).  $T$  is also used time-dependent as  $T(t)$ .

**Herd immunity:** The herd immunity  $H(t)$  for competitive serotypes is defined based on single-pathogen-case:

**Herd immunity in case of competitive serotypes indicates a benefit for unvaccinated persons by occurrence of vaccinated persons related to the serotype(s) included in the vaccination.**

This means that the herd immunity is an effect that only depends in serotype V. It can be described by equation 7.1.

$$H(t) = \frac{T_{N,base,V}(t) - T_{N,test,V}(t)}{T_{N,base,V}(t)} \quad (7.1)$$

**Serotype shift:** The definition of serotype shift refers to the remarks in chapter 7.2.

**Serotype shift, in case of competitive serotypes, indicates a disadvantage by serotypes that are not included in a vaccination. It potentially occurs when people are vaccinated and affects the whole population.**

Equation 7.2 shows the calculation of serotype shift  $S(t)$ .

$$S(t) = \frac{T_{A,test,N}(t) - T_{A,base,N}(t)}{T_{A,base,N}(t)} \quad (7.2)$$

**Alternative consideration of serotype shift:** As said before, it might be necessary to differentiate this effect for vaccinated and unvaccinated people.

In this context a question comes up: Whom should the vaccinated population be compared with? If there are no vaccinated people in the base scenario, the only possibility is to compare them to the unvaccinated people. If there are vaccinated people in the base scenario, they could be compared to them.

In other words, this means that if there are no vaccinated people in the base scenario, they would expect the same disadvantage by serotype N after the vaccination. If there are vaccinated people in the base scenario, then unvaccinated people would expect the same disadvantage upon vaccination by serotype N as already vaccinated people.

This leads to the following equation 7.3:

$$\begin{aligned} S_N(t) &= \frac{T_{N,test,N}(t) - T_{N,base,N}(t)}{T_{N,base,N}(t)} \\ S_V(t) &= \frac{T_{V,test,N}(t) - T_{\{N|V\},base,N}(t)}{T_{\{N|V\},base,N}(t)} \end{aligned} \quad (7.3)$$

**General herd immunity:** Another important consideration is the change of the total disadvantage by a vaccination, independent of the serotype. This effect affects the whole population; it contains the direct reduction of serotype V by the vaccinations, the herd immunity and the serotype shift. It seems practicable to name this effect “general herd immunity” and abbreviate it by  $GH(t)$ . This quantity is important if the characterization of the whole disadvantage does not need any differentiation between serotypes,



even though the vaccination only works against one serotype. Equation 7.4 shows the calculation of this effect.

$$GH(t) = \frac{T_{A,base,V\&N}(t) - T_{A,test,V\&N}(t)}{T_{A,base,V\&N}(t)} \quad (7.4)$$

The general herd immunity can also be calculated with already known quantities if the disadvantage of any serotype equals the sum of the disadvantages of serotypes N and V. This means that  $T_{x,y,V\&N} = T_{x,y,V} + T_{x,y,N}$ . The equations 7.5 and 7.6 show these relations by differentiation between the two serotypes. All quantities in the formulas are time-dependent but  $(t)$  has been removed due to better readability.

$$\begin{aligned} GH &= \frac{[T_{N,base,V} + T_{A,base,N}] - [T_{N,test,V} + T_{A,test,N}]}{T_{N,base,V} + T_{A,base,N}} \\ &= \frac{[T_{N,base,V} + T_{A,base,N}] - [T_{N,base,V} \cdot (1 - H) + T_{A,base,N} \cdot (1 + S)]}{T_{N,base,V} + T_{A,base,N}} \\ &= \frac{T_{N,base,V} \cdot H - T_{A,base,N} \cdot S}{T_{N,base,V} + T_{A,base,N}} \end{aligned} \quad (7.5)$$

$$\begin{aligned} GH &= \frac{[T_{N,base,V} + T_{N,base,N} + T_{V,base,N}] - [T_{N,test,V} + T_{N,test,N} + T_{V,test,N}]}{T_{N,base,V} + T_{N,base,N} + T_{V,base,N}} \\ &= \frac{T_{N,base,V} + T_{N,base,N} + T_{V,base,N}}{T_{N,base,V} + T_{N,base,N} + T_{V,base,N}} - \\ &\quad - \frac{T_{N,base,V} \cdot (1 - H) + T_{N,base,N} \cdot (1 + S_N) + T_{V,base,N} \cdot (1 + S_V)}{T_{N,base,V} + T_{N,base,N} + T_{V,base,N}} \\ &= \frac{T_{N,base,V} \cdot H - T_{N,base,N} \cdot S_N - T_{V,base,N} \cdot S_V}{T_{N,base,V} + T_{N,base,N} + T_{V,base,N}} \end{aligned} \quad (7.6)$$

**Complete herd immunity:** Complete herd immunity in case of competitive serotypes indicates a state where Serotype V cannot survive any more and becomes extinct. This definition is consistent to the one for single pathogens because this state is characterized by the situation that serotype V does not cause any disadvantage any more. Hence, the herd immunity reaches its maximum.

**Generalized calculations of herd immunity and serotype shift:** So far the definitions have always been limited to the whole population or the vaccinated or unvaccinated persons. However, there might be interest in particular groups of the population and the effects of vaccinations on them, no matter if they are vaccinated or not. Especially for non-homogeneous populations such calculations might be useful. The definitions generally allow to extend the calculations on any kind of such groups or population parts. One

just has to keep in mind that some people do not necessarily exist in the base scenario if the group is defined by their vaccination state. Hence, a useful reference group has to be defined then.

## 7.4 Definition of the disadvantage

Consideration and calculation of the disadvantage works the same way like for single serotypes, as described in chapter 4.4.2. This means that the disadvantage primarily corresponds to the number of new infections because they are the immediate effect of a disease. Carrier rates or sickness cases are examples of alternative quantifications. The temporal component can be used in three ways that have to be selected depending on the actual disease: Infections in time  $t$ , Infections in the last  $x$  time step before time  $t$  and all infections since simulation start until time  $t$ .

Again it is important to normalize the disadvantage which means to divide by the size of the considered group. Usually there are less unvaccinated people in the test scenario than in the base scenario. Hence, a comparison of absolute infection numbers is neither correct nor meaningful.

## 7.5 Herd immunity and serotype shift as an explaining quantity

The goal of this chapter is to examine herd immunity and serotype shift with the agent based model and interpret them.

### Model setup:

- The model and the parameters settings are the same as in chapter 6.5. However, only the case with different infection probabilities ( $p_1 = 0.022$ ,  $p_2 = 0.02$ ) is used here as an exemplary simulation. Extensive parameter variations are shown in the following chapters.
- All simulations use approach 2 - competition with mutual susceptibility - because approach 1 does not allow to produce stable equilibriums for SIS epidemics.
- **Scenarios:**
  - **Base scenario:** In the base scenario, no people are vaccinated.
  - **Test scenario:** The test scenario is calculated with a vaccination rate of 20% against serotype 1. According to the different infection probabilities, this is the stronger serotype. People are vaccinated at the simulation start and they are chosen randomly among the whole population.

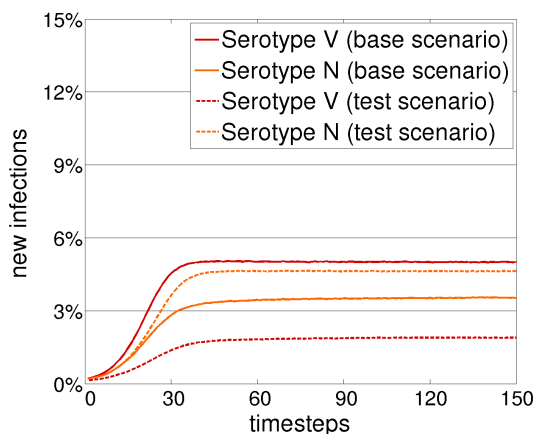
- **Disadvantage:** The disadvantage is calculated by three methods: Current new infections, new infections since simulation start and current carrier rate. Hence, three different results for herd immunity and serotype shift are calculated.

### 7.5.1 Herd immunity and serotype shift in SIS models

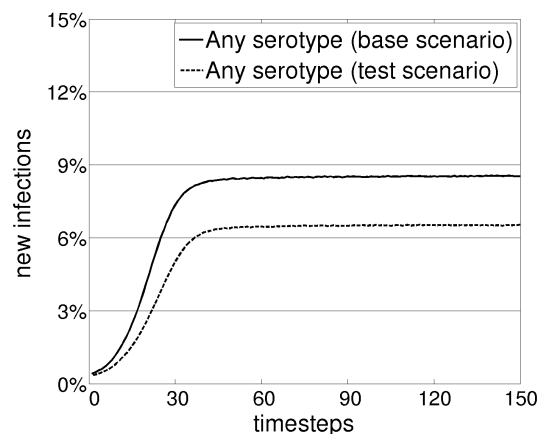
The results of the simulations are shown as new infection curves (figures 7.2 and 7.3) and carrier rates (figures 7.4 and 7.5). Both simulations show the typical SIS behavior which results in an equilibrium after a short adaption phase. It is obvious that both suspected effects occur: The vaccinations lead to a lower level of serotype V and a higher level of serotype N. This means that serotype N indeed makes use of the fewer competition and infects more people than before. Altogether, the vaccination has a positive effect as the serotype-independent infection curves and carrier rates show.

These effects are also proved by herd immunity and serotype shift (figures 7.6, 7.7 and 7.8). All three calculation methods of the herd immunity show an increase from zero in the beginning to peak right before time 30 and then a drop to almost 0.6. This means that the unvaccinated population can expect a benefit concerning infections with serotype V by 60%. On the other, hand there is a strong serotype shift. If it is calculated by the new infections, it is about 0.3 in the equilibrium but for the carrier rate it is about 0.7. This discrepancy results from the fact that in every time step several infected people get infected with the other serotype. Therefore, the number of new infections and the carrier rate does not correspond to each other.

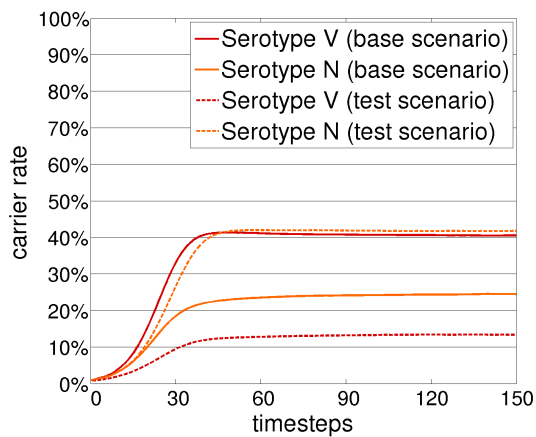
The general herd immunity (figures 7.6, 7.7 and 7.8) reaches about 0.2 in the equilibrium. Considering the 20% vaccination rate, this means that altogether herd immunity and serotype shift almost neutralize each other and the positive effect for the population is more or less the vaccination rate. Later, the parameter studies will show that this does not always apply and different situation can lead to quite different epidemic dynamics.



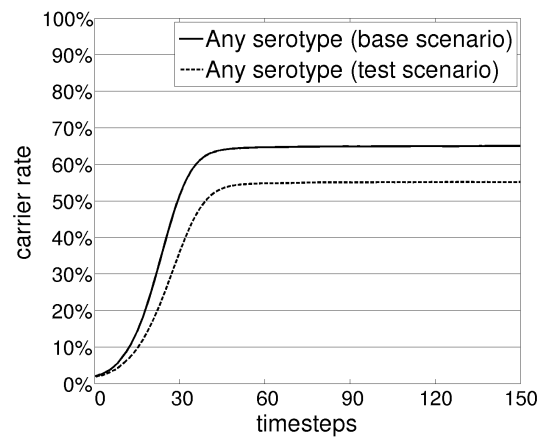
**Figure 7.2:** New infections by serotype in the SIS simulations



**Figure 7.3:** Total new infections in the SIS simulations



**Figure 7.4:** Carrier rates by serotype in the SIS simulations



**Figure 7.5:** Total carrier rates in the SIS simulations

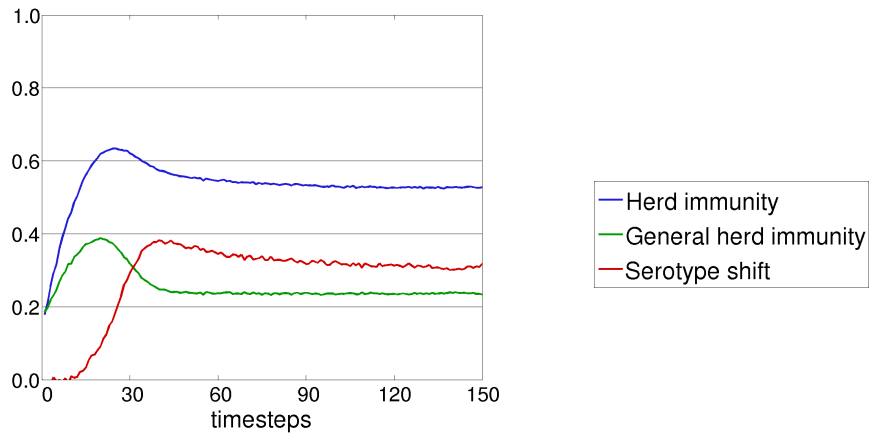
**Variation:** An alternative vaccination strategy for SIS epidemics is to wait, until the base scenario has reached its equilibrium and vaccinate then. The equilibrium after the vaccination is the same since it does not depend on the initial values, but the adaptation phases show an interesting behavior. While the total number of infections reaches a new level very quickly after vaccination, the redistribution of the living space is a process that takes much longer. The figures 7.9 and 7.10 show the new infections and the carriers, while the figure 7.11 shows the herd immunity, the serotype shift and the general herd immunity (all calculated upon new infections).

## 7.5.2 Herd immunity and serotype shift in SIR models

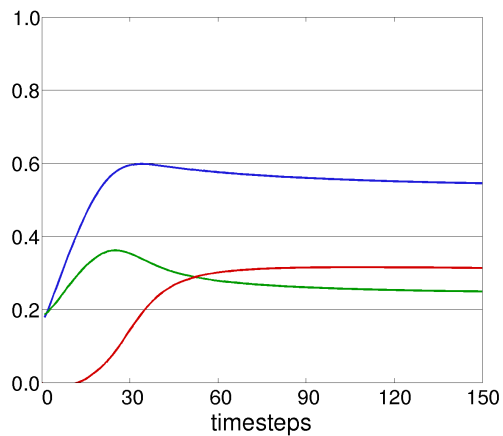
The tests in the SIR model are performed analogously to the SIS model. The figures 7.12 and 7.13 show the new infections, while the figures 7.14 and 7.15 show the carrier rates. It is obvious that the vaccinations strongly decreases the prevalence of serotype V while an increased prevalence of serotype N can be observed.

If herd immunity and serotype shift are calculated upon current infections or carrier rates, they raise the same issues as in the single-pathogen-case (see figures 7.16 and 7.18). The herd immunity and the general herd immunity are strongly positive at first, but after about 50 time steps they are becoming negative. The problem is that later in the simulation, the epidemic is already dying out and only a few more people get infected. Since the epidemic is distorted to the right by the vaccination, more infections happen in the test case. By comparing very small numbers for the disadvantage, the herd immunity is represented by large negative numbers. This means that there might be 10 or 20 times more infections with serotype N in the test case after time 70, but the amount of people after this time is almost negligible compared to the infections before.

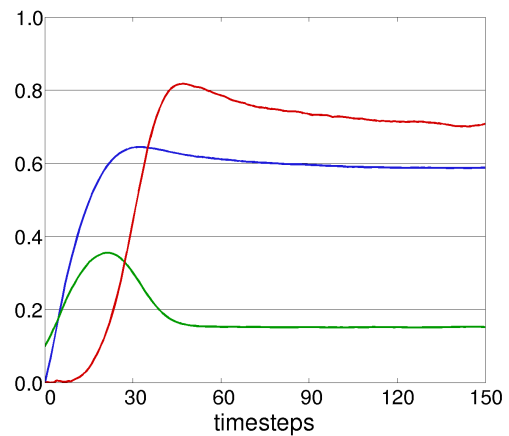
The calculations based on the infections since simulation start are more meaningful as shown in figure 7.17 and can therefore be interpreted well. Both herd immunity and



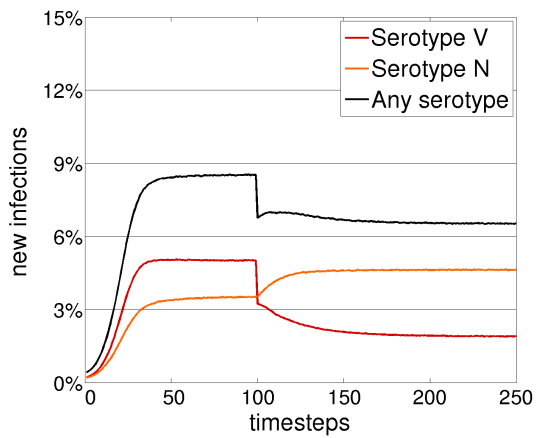
**Figure 7.6:** Epidemic effects in the SIS simulation, calculated upon current new infections



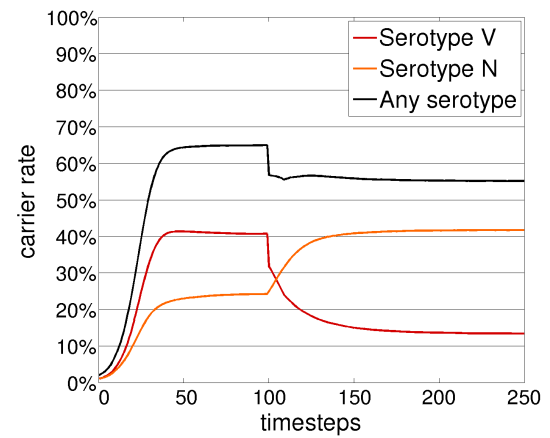
**Figure 7.7:** Epidemic effects in the SIS simulation, calculated upon new infections since start



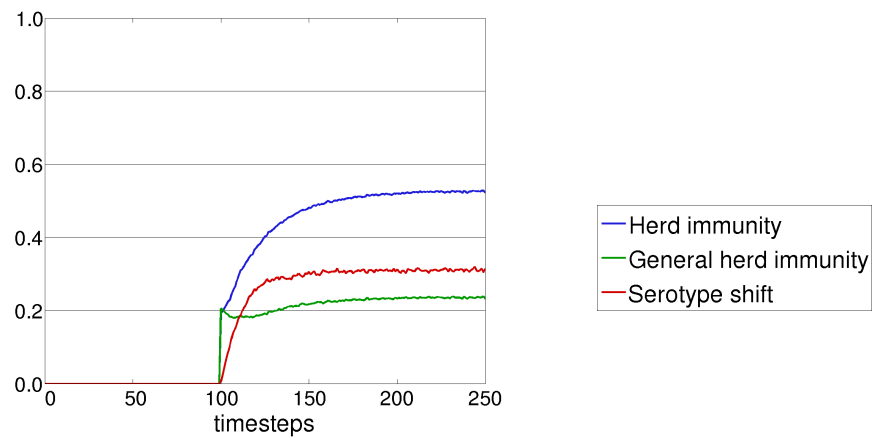
**Figure 7.8:** Epidemic effects in the SIS simulation, calculated upon the carrier rate



**Figure 7.9:** New infections in the SIS simulation with vaccination after 100 time steps



**Figure 7.10:** Carrier rates in the SIS simulation with vaccination after 100 time steps

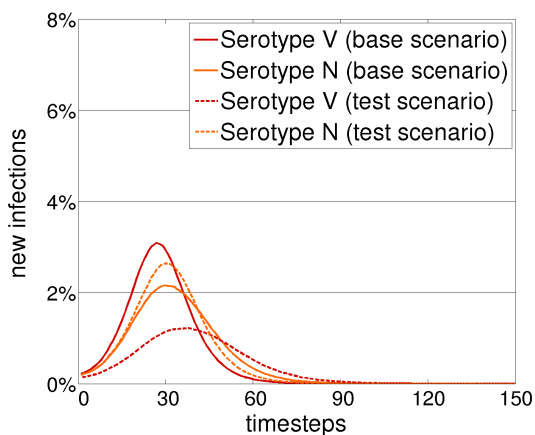


**Figure 7.11:** Epidemic effects in the SIS simulation with vaccination after 100 time steps, calculated upon current new infections

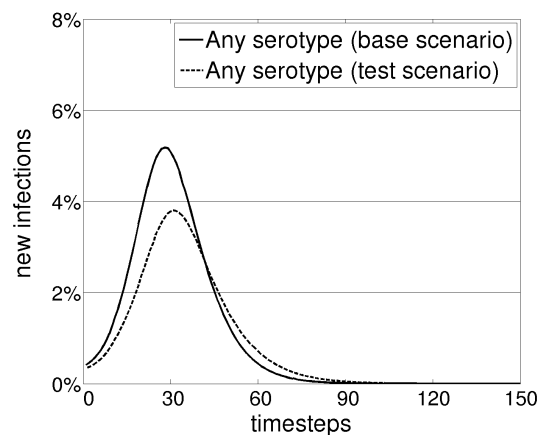
## 7.5. HERD IMMUNITY AND SEROTYPE SHIFT AS AN EXPLAINING QUANTITY

---

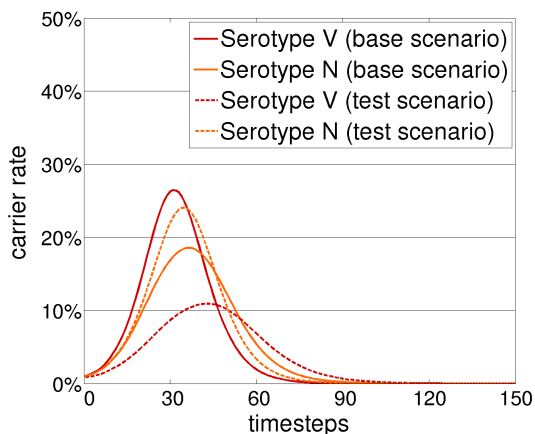
general herd immunity reach a level slightly below 0.2 at the end of the epidemic. There is a strong serotype shift in the beginning but in the end it is very weak. This means that the disadvantage by serotype N of the whole epidemic is almost the same in both scenarios. It just has different temporal progressions. Herd immunity is clearly prevalent and represent a benefit of almost 20% concerning serotype V for the unvaccinated population. Altogether the general herd immunity shows a benefit of almost 20% for the whole population for all serotypes. In other words, thanks to the herd immunity and low serotype shift, the 20% vaccination rate against one serotype only reduces the total disadvantage by both serotypes by 20%.



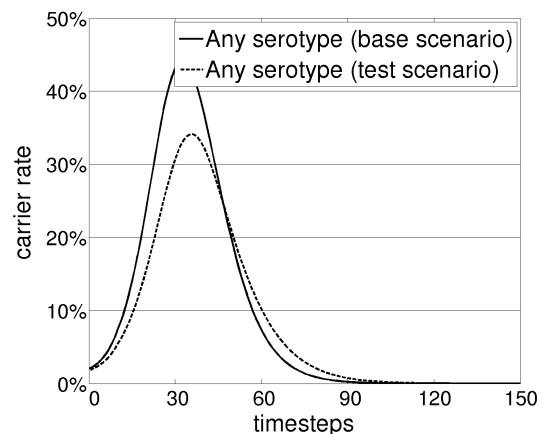
**Figure 7.12:** New infections by serotype in the SIR simulation



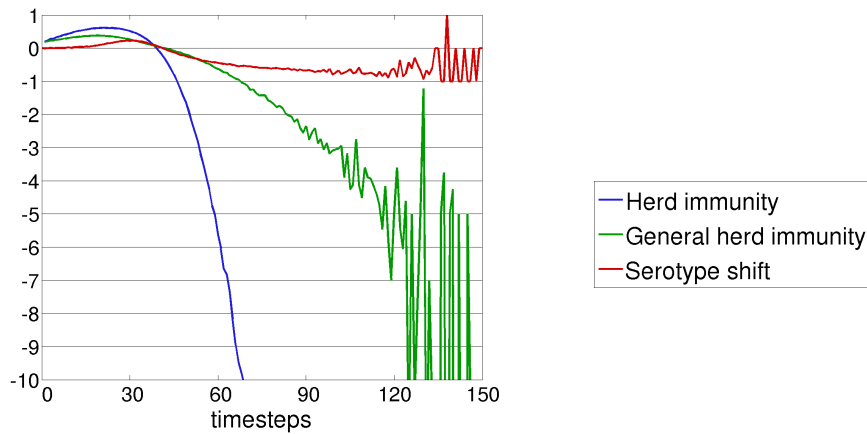
**Figure 7.13:** Total new infections in the SIR simulation



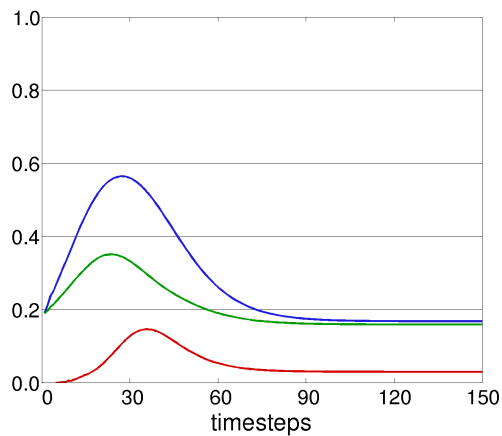
**Figure 7.14:** Carrier rates by serotype in the SIR simulation



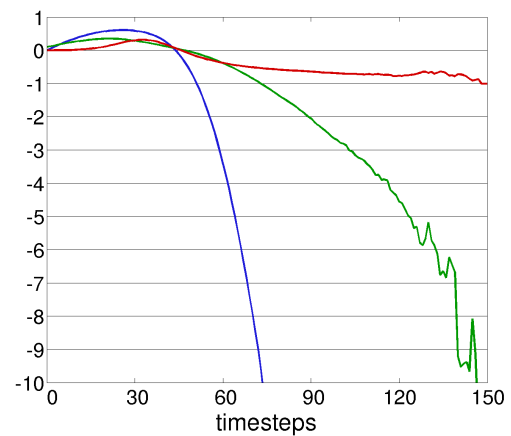
**Figure 7.15:** Total carrier rates in the SIR simulation



**Figure 7.16:** Epidemic effects in the SIR simulation, calculated upon current new infections



**Figure 7.17:** Epidemic effects in the SIR simulation, calculated upon new infections since start



**Figure 7.18:** Epidemic effects in the SIR simulation, calculated upon the carrier rate



## 7.6 Parameter studies for herd immunity and serotype shift

Based on the used and studied simulation which is defined in chapter 6.5, the infection probability and the recovery time of Serotype V are varied. Due to the relation between infection probability and contact number, it is not necessary to vary contacts as well. Both parameters are adjusted by 10%, 20% and 30% lower and higher; table 7.1 shows the concrete settings. Variations of Serotype N are not performed because the current test series already includes both weaker and stronger serotypes V. Combinations of parameter variations are also not examined which means that each simulation differs in exactly one parameter from the base setting.

**Table 7.1:** Settings of the parameter variation for competitive serotypes

Variation of serotype V	-30%	-20%	-10%	base	+10%	+20%	+30%
Infection probability	0.014	0.016	0.018	0.02	0.022	0.024	0.026
Recovery time	7	8	9	10	11	12	13

### 7.6.1 Parameter variations of the SIS model

The parameter variations are calculated as defined above. The results of the SIS model are the new infections and the carrier rates of the equilibrium.

**Variation of the infection probability:** The figures 7.19 and 7.20 show the results of the variation of the infection probability. It is clear that a higher infection probability causes more infections with serotype V. However, due to the competition with serotype N the total carrier does not reach a level above 60% and the total infection numbers even decrease for a too strong serotype V. Herd immunity, serotype shift and general herd immunity are shown in figures 7.21, 7.22 and 7.23.

**Variation of the recovery time:** The results are shown in figures 7.24 and 7.25 while the epidemic effects in figures 7.26, 7.27 and 7.28. The impact of parameter changes is similar to the infection probability test series but is quantitatively less strong. Presumably, quantitative results that are similar to the first test series can be generated by a stronger variation of the recovery time.

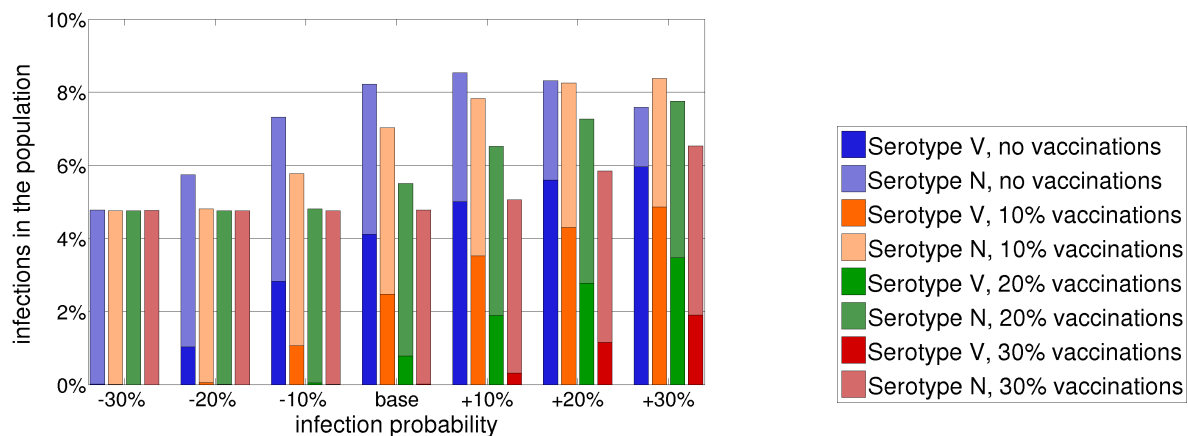
**Interpretation of the results:** Complete herd immunity can be observed in a few cases where serotype V becomes extinct due to the vaccination. In this cases, herd immunity

## CHAPTER 7. HERD IMMUNITY AND SEROTYPE SHIFT IN CASE OF TWO COMPETITIVE SEROTYPES

equals 1. The herd immunity is mostly high but gets weaker as stronger serotype V is. A higher vaccination rate results in a stronger herd immunity.

Serotype shift shows an interesting behavior. In most cases, the serotype shift is not very strong, even though it is growing with the strength of serotype V, which causes serotype N to be weak in the simulations without vaccinations. It also grows with the vaccination rate. But setting the infection probability +30%, the serotype shift is even higher than 1. This means that serotype N is on a very low level due to the competition. Hence it can become more than twice as strong by vaccination against serotype V.

The general herd immunity, which indicates the total effect of the vaccination, naturally increases as the vaccination rate increases. However, it has its peak around the base simulation which means that the overall outcome of the vaccinations reach its maximum when both serotypes are equipollent or when the vaccinated serotype is slightly stronger. Again for +30% infection probability, an interesting effect occurs. In this case, the general herd immunity is negative which means that the serotype shift due to a few vaccinations of 10% or 20% is stronger than the herd immunity, so there is no benefit by the vaccination at all. This effect can be explained that Serotype N spreads so much stronger in these cases that it cannot be compensated by the vaccination rate and the herd immunity of serotype V. However, 30% vaccination rate already causes a strong enough herd immunity which leads to a positive effect again.

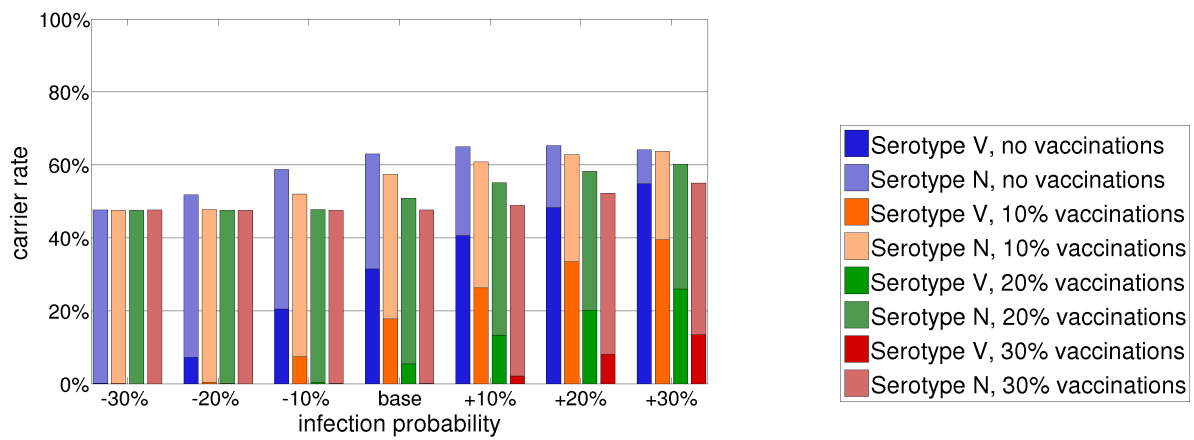


**Figure 7.19:** New infections per time step (variation of the infection probability in the SIS model)

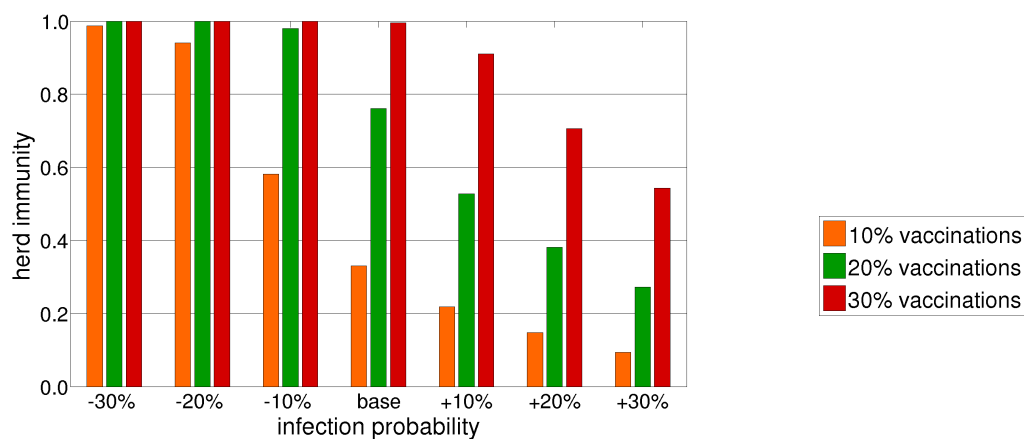
### 7.6.2 Parameter variations of the SIR model

In the SIR model, the results are represented by the total number of infections during the whole epidemic. The temporal propagation is not considered. This corresponds to the calculation of the epidemic effects which are based on the infections since start. They are represented by their value at the end of the simulation, where the disease is already extinct, so they cover the whole epidemic.

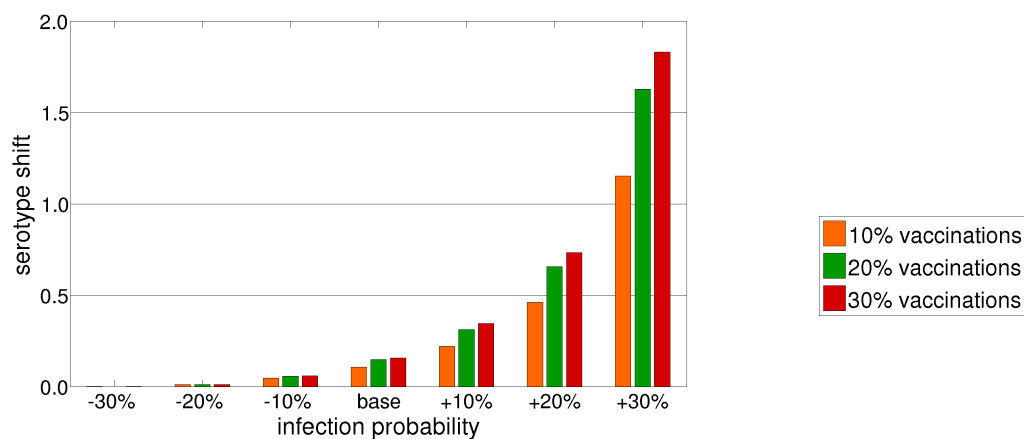
## 7.6. PARAMETER STUDIES FOR HERD IMMUNITY AND SEROTYPE SHIFT



**Figure 7.20:** Carrier rates (variation of the infection probability in the SIS model)

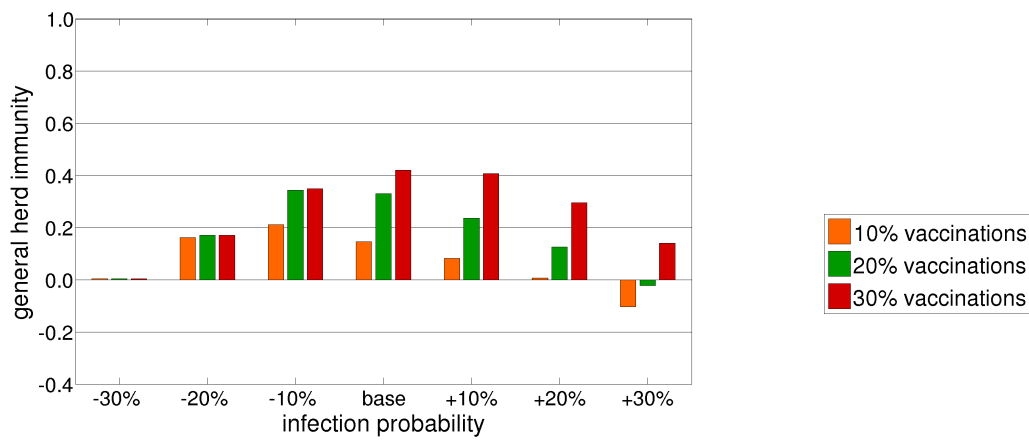


**Figure 7.21:** The herd immunity (variation of the infection probability in the SIS model)

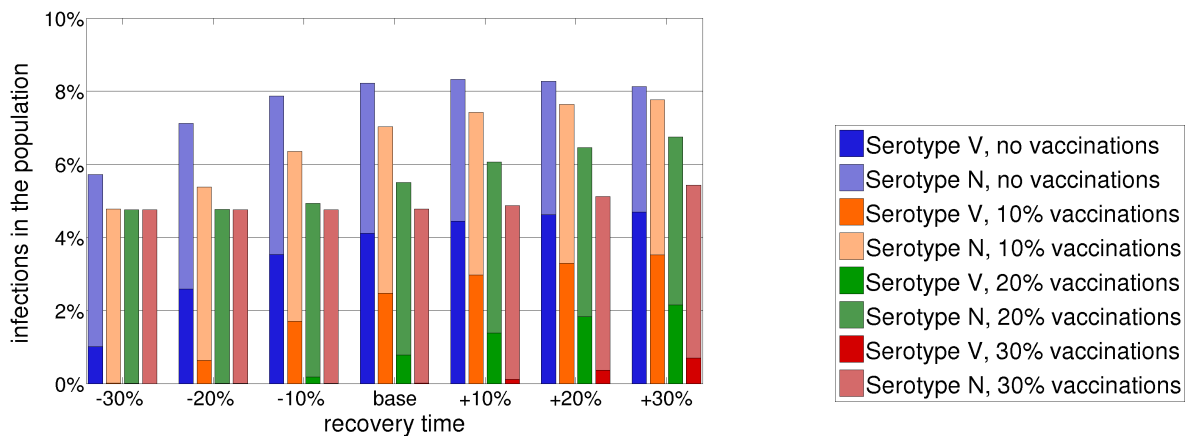


**Figure 7.22:** The serotype shift (variation of the infection probability in the SIS model)

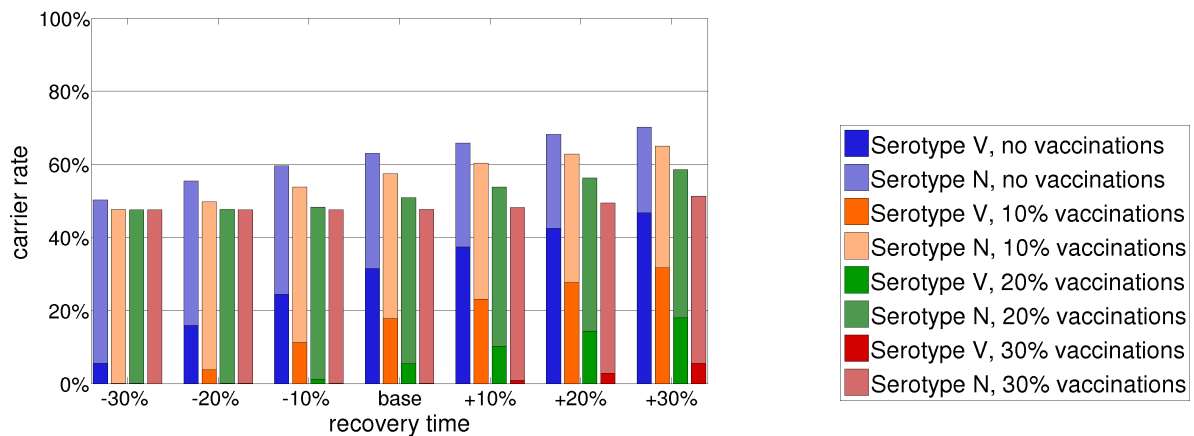
CHAPTER 7. HERD IMMUNITY AND SEROTYPE SHIFT IN CASE OF TWO COMPETITIVE SEROTYPES



**Figure 7.23:** The general herd immunity (variation of the infection probability in the SIS model)

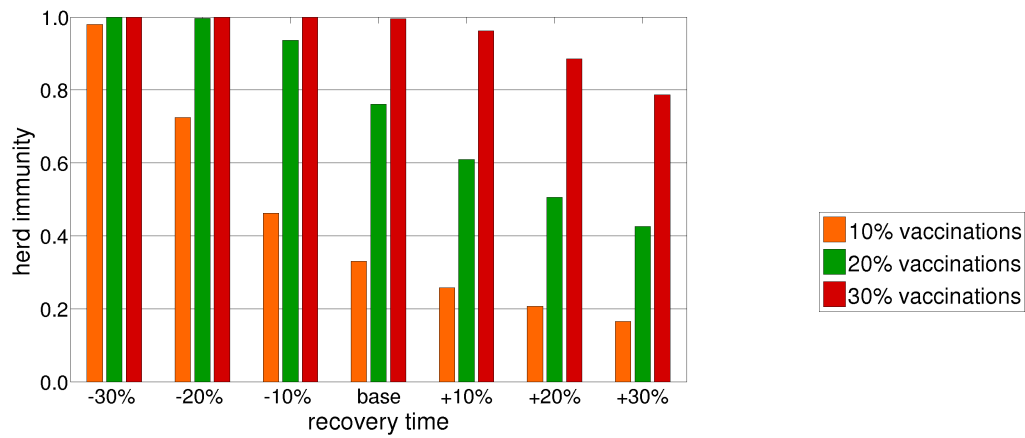


**Figure 7.24:** New infections per time step (variation of the recovery time in the SIS model)

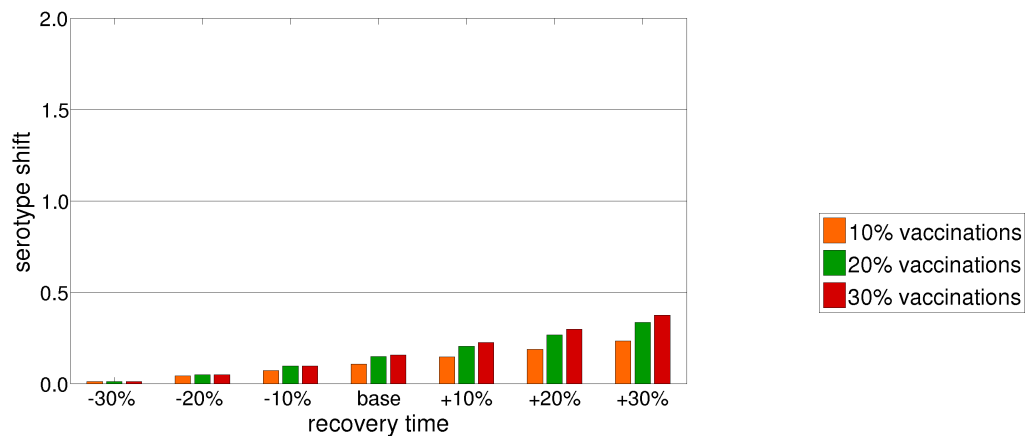


**Figure 7.25:** Carrier rates (variation of the recovery time in the SIS model)

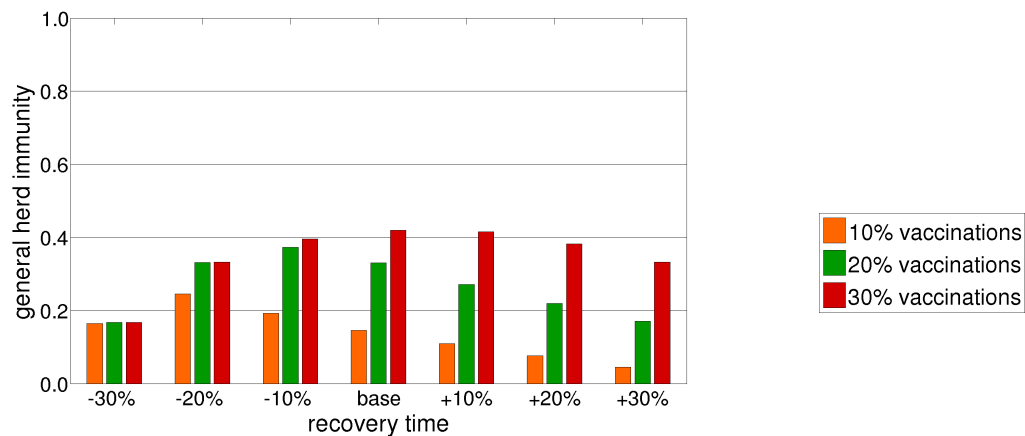
## 7.6. PARAMETER STUDIES FOR HERD IMMUNITY AND SEROTYPE SHIFT



**Figure 7.26:** The herd immunity (variation of the recovery time in the SIS model)



**Figure 7.27:** The serotype shift (variation of the recovery time in the SIS model)



**Figure 7.28:** The general herd immunity (variation of the recovery time in the SIS model)

**Variation of the infection probability:** The figure 7.29 shows the results of the variation of the infection probability. Obviously, a higher infection probability causes a higher number of infected people. The total infections vary between 80% and 160%. The number of infections can be higher than 100% because each person can be infected twice, with serotype V and N, therefore the theoretical maximum is 200%. Vaccinations do what they are ought to do and decrease the infection numbers.

**Variation of the recovery time:** The results of the variation of the recovery time are shown in figure 7.33. They are very similar to the results of the variation of the infection probability.

**Interpretation of the results:** Herd immunity, serotype shift and general herd immunity are shown in figures 7.30, 7.31 and 7.32 respectively 7.34, 7.35 and 7.36.

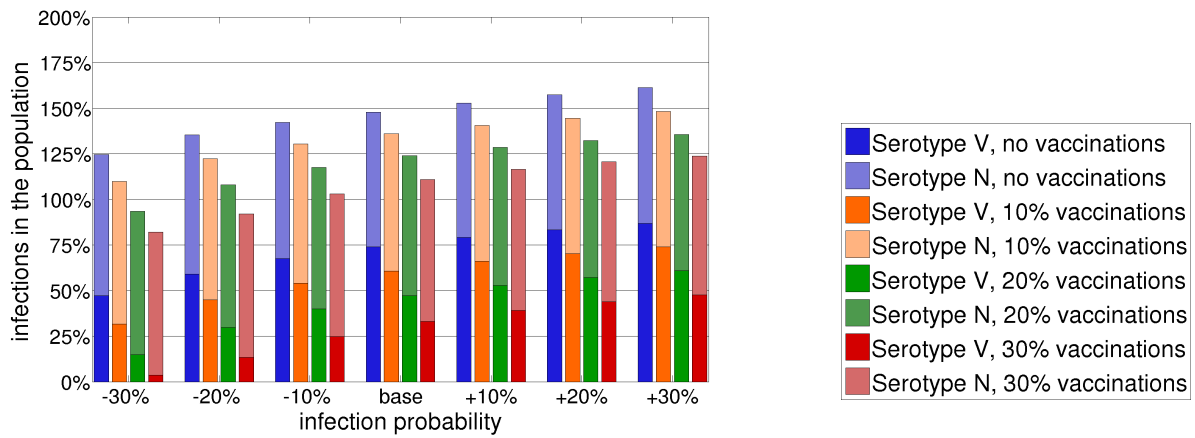
Complete herd immunity cannot be observed, only in case of -30% infection probability or recovery time, 30% vaccinations can almost prevent an outbreak of serotype V. The herd immunity is very high for a weak serotype V but it is rapidly decreasing to lower levels as serotype V gets weaker. Altogether, this means that the unvaccinated population can expect a benefit concerning serotype V. This benefit is growing with a higher vaccination rate and with a weaker serotype V.

The serotype shift is on very low level, always below 0.08 which is much lower than for SIS epidemics. For different infection probabilities, it has a maximum for equally strong serotypes and it is lower for other settings for serotype V. In one case, it is even negative which means that serotype N gets weaker due to the vaccination. The serotype shift grows for longer recovery time but at some point, around the base setting, it reaches a maximum. Generally, we can state that the impact of the vaccination on the number of infections with serotype N is very low, typically Serotype N infects 1% to 5% more people due to vaccinations against serotype V.

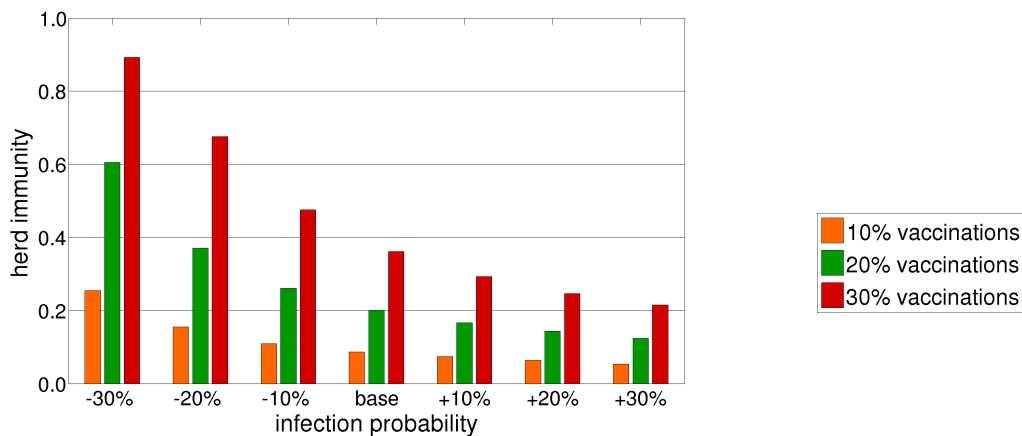
The general herd immunity is almost the same for both test series. For a weak serotype V, it almost corresponds to the vaccination rate which means that the reduction of all infections equals the part of the population that gets vaccinated against serotype V. If serotype V gets stronger, then herd immunity decreases and serotype shift increases, hence more infections happen. The general herd immunity shows this fact clearly because it is decreasing for higher infection probabilities and recovery times. In other words, the reduction of total infections is as lower as the vaccine-serotype gets stronger.

## 7.7 Conclusions

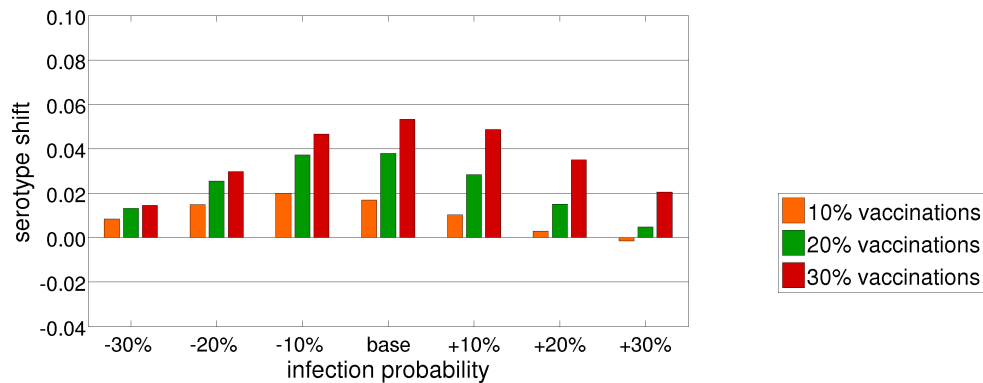
In this chapter, three terms for effects of epidemics with competitive serotypes have been defined: The herd immunity shows the benefit for the unvaccinated population concerning the vaccine-serotype, the serotype shift shows the disadvantage for the whole population



**Figure 7.29:** Total number of infections (variation of the infection probability in the SIR model)

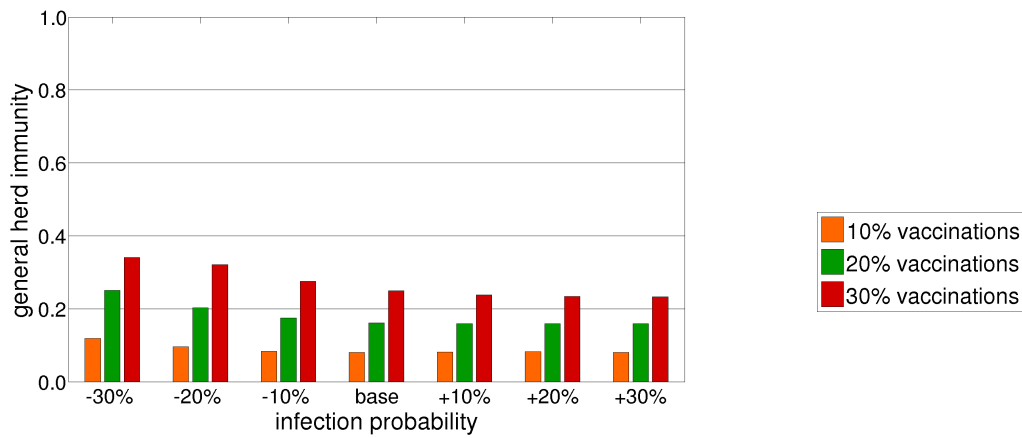


**Figure 7.30:** The herd immunity (variation of the infection probability in the SIR model)

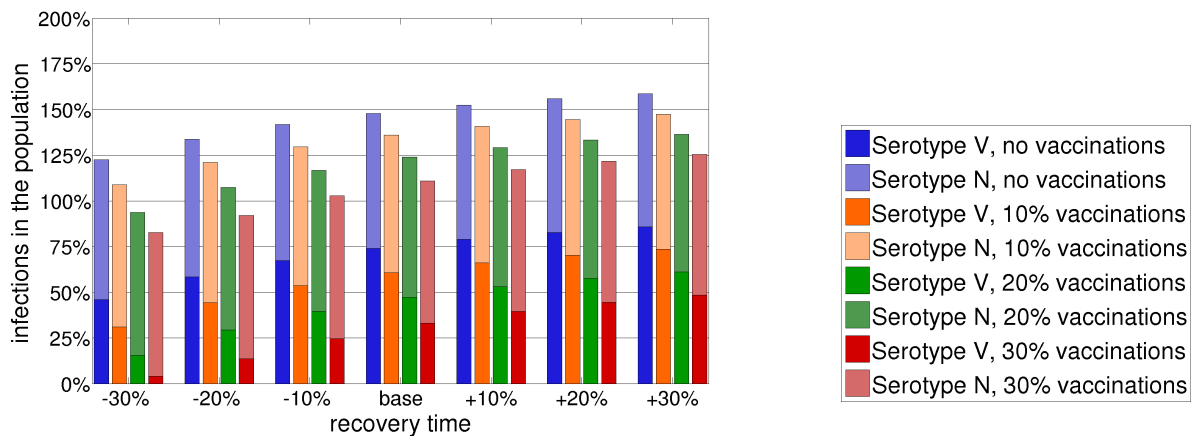


**Figure 7.31:** The serotype shift (variation of the infection probability in the SIR model)

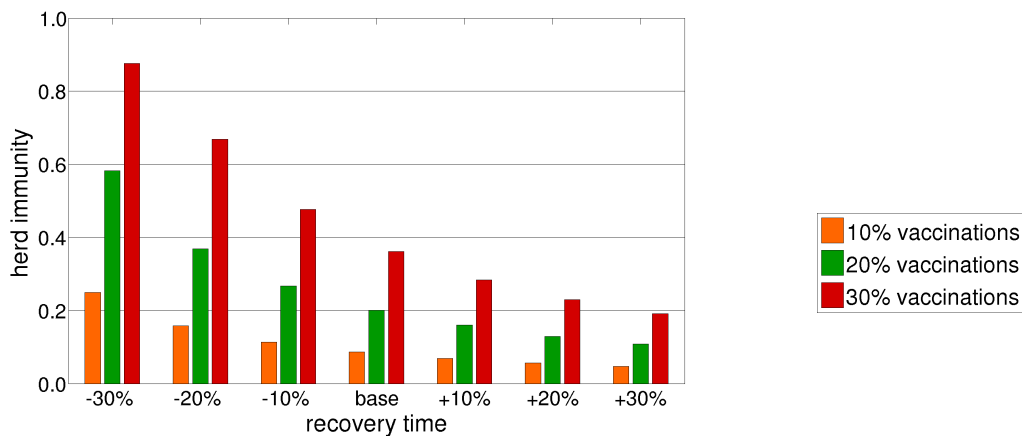
CHAPTER 7. HERD IMMUNITY AND SEROTYPE SHIFT IN CASE OF TWO COMPETITIVE SEROTYPES



**Figure 7.32:** The general herd immunity (variation of the infection probability in the SIR model)

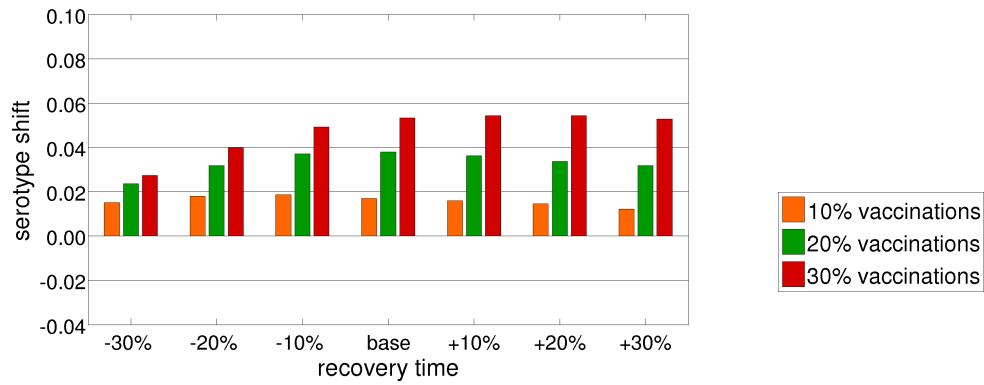


**Figure 7.33:** Total number of infections (variation of the recovery time in the SIR model)

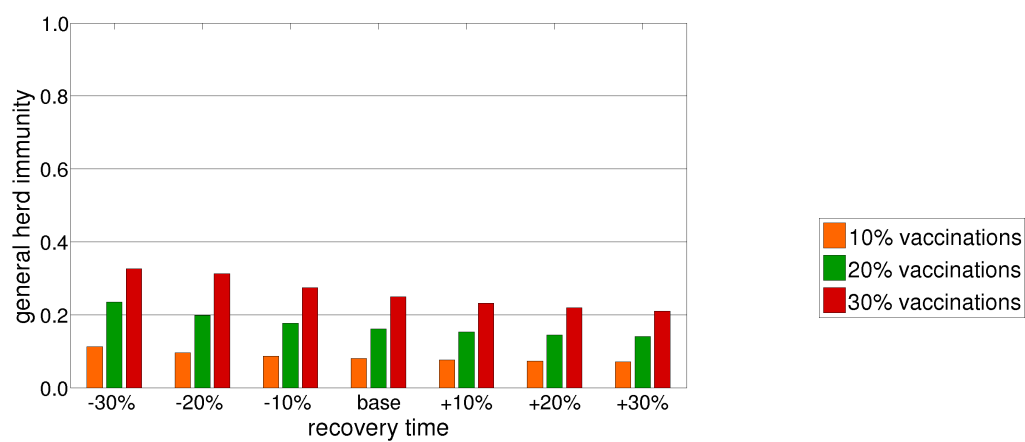


**Figure 7.34:** The herd immunity (variation of the recovery time in the SIR model)





**Figure 7.35:** The serotype shift (variation of the recovery time in the SIR model)



**Figure 7.36:** The general herd immunity (variation of the recovery time in the SIR model)

## CHAPTER 7. HERD IMMUNITY AND SEROTYPE SHIFT IN CASE OF TWO COMPETITIVE SEROTYPES

---

concerning the other serotype and the general herd immunity shows the total benefit of vaccinations. Again, these effects rely on a correct calculation of the disadvantage by the disease. If the time component is chosen appropriately then these three effects can describe the impact of a vaccination strategy very well. For SIS epidemics, they are best interpreted by the current infections in the equilibrium while SIR epidemics require the sum of infections since start for every time point for useful and descriptive results.

# Chapter 8

## Comparison of Markov and agent based models for epidemic simulations with two competitive serotypes

### 8.1 Overview

#### 8.1.1 Introduction

A comparison between Markov models and agent based models is also possible for epidemics with two competitive serotypes. The procedure is the same as in chapter 5:

1. Construction of an agent based and a Markov model for simulation of an epidemic.
2. Parameterization of both models so that their results correspond to each other for simulations without vaccinations.
3. Implementation of different vaccination strategies on the population in both models and comparison of the results.
4. Presentation of the diverging results and explanation the differences from a mathematical point of view.

The examined examples are also similar to the single-pathogen-case since they are easy to handle and are proven to yield a descriptive comparison.

1. Comparison of equilibriums in an SIS model.
2. Comparison of the temporal propagation of an SIR model.
3. Comparison of infection numbers of a complete season in an SIR model.

### 8.1.2 Considerations about comparability of agent based and Markov models with two serotypes

The models with a single pathogen provided direct comparability. Two serotypes lead to much more complex models with more states and more possible transitions. Hence, a state in the Markov model for competitive serotypes often depends on several transition probabilities. These probabilities can always be measured in the agent based model and allow to reproduce its results with the Markov model. But they cannot be adjusted easily to obtain a desired output. Hence, calibration of the Markov model to reach the results of the agent based model is a multi-dimensional optimization problem and would not work well in this case. However, as long as direct parameterization is possible, calibration is not needed. Additionally, it is required that both models produce results of the same structure to provide comparability. The results of the Markov model are not necessarily the states itself. Often they are derived from the states and from state changes.

Herd immunity and serotype shift are considered in a way that they are applied on the results of the Markov-Model, like for models of single pathogens. This means that if new infections or the carrier rate are the results of the Markov model, herd immunity and serotype shift are calculated based on the same quantities.

An important question is how the vaccinated part of the population in the Markov model should be handled because there are no vaccinated people in the base case of the used simulations. Separation of the vaccinated and the unvaccinated population parts follows the latest standard which is also suggested by the ISPOR-SMDM Joint Modeling Good Research Practices Task Force [51]. Hence, the vaccinated part of the population must be parameterized in a way that new infections and carriers with serotype N are the same as in the unvaccinated population. The same problem already occurs for calculation of herd immunity and serotype shift. The proposed solution in chapter 7.3 corresponds to the strategy here. This is important to use and to apply herd immunity and serotype shift correctly. If there are already vaccinated people in the base scenario, then their parameters can be measured directly.

### 8.1.3 The models

**The agent based model:** The agent based model is the one that has been used already in the previous chapters 6 and 7. To reach equilibriums the competition is again realized by approach 2, which means mutual susceptibility. Hence, the definition of the used agent based model originates from chapter 6.6 based on the definition in chapter 6.5.

The parameterization of the agent based model is the same as set in chapter 6.5, with unequal serotypes by different infection probabilities and vaccinations against the stronger serotype. This simulation has been examined already for stability (chapters 6.6), vaccinations, herd immunity and serotype shift (chapter 7).

**The settings for the agent based model:**

- Number of persons:  $N = 5$  million
- Contacts per persons per time step:  $k = 10$
- Disease type: The model can be used for both SIS and SIR diseases.
- Start values: In the initial state 1% of the population is infected with serotype 1 and serotype 2 each and 98% are susceptible.
- Infection probability:  $p_1 = 0.022$  for serotype 1 and  $p_2 = 0.02$  for serotype 2
- Recovery time:  $d = 10$  time steps
- Competition: Approach 2 (mutual susceptibility)
- Base scenario: No vaccinations
- Test scenario: Three test scenarios are simulated with 10%, 20% and 30% vaccination rates against serotype 1.

**The Markov model:** The Markov model cannot be built in general. Therefore, it is described separately for each test case. However, the population in all examples is split into unvaccinated and vaccinated persons who are simulated with different models.

## 8.2 Comparison of equilibriums in an SIS epidemic

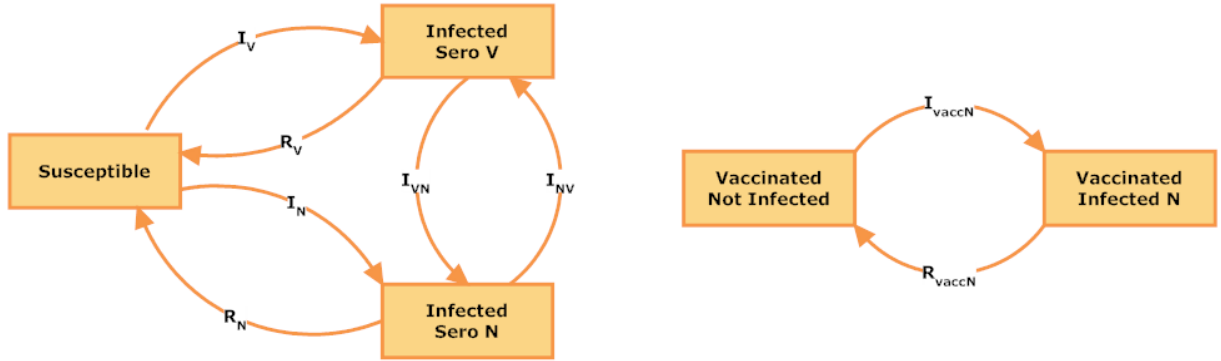
### 8.2.1 Construction of the models

**The agent based model:** The agent based model calculates an SIS epidemic and is defined in chapter 8.1.3.

**The Markov model:** The Markov model is much more complex than in case of single pathogens. Figure 8.1 shows the state transition diagrams, on the left for unvaccinated people and on the right for vaccinated people. Figures 8.2 visualizes the tree of the whole model.

**Unvaccinated persons:** Mathematically, a distribution of this system in time  $t$  is represented by a 3-dimensional vector  $\underline{a}_t = (a_{1,t}, a_{2,t}, a_{3,t})$  following the states  $s_1 =$ “susceptible”,  $s_2 =$ “infected with serotype V” and  $s_3 =$ “infected with serotype N”. Hence, the transition matrix looks like this:

$$M = \begin{pmatrix} 1 - I_V - I_N & I_V & I_N \\ R_V & 1 - R_V - I_{VN} & I_{NV} \\ R_N & I_{VN} & 1 - R_N - I_{NV} \end{pmatrix}$$



**Figure 8.1:** The state transition diagram of a Markov model in case of two competitive serotypes.

**Left:** For unvaccinated persons. **Right:** For vaccinated persons.

**Vaccinated persons:** The distribution of the system in time  $t$  is represented by a 2-dimensional vector  $\tilde{\underline{a}}_t = (\tilde{a}_{1,t}, \tilde{a}_{2,t})$  following the states  $\tilde{s}_1 =$ “vaccinated and susceptible” and  $\tilde{s}_2 =$ “vaccinated and infected with serotype N”. The transition matrix looks like this:

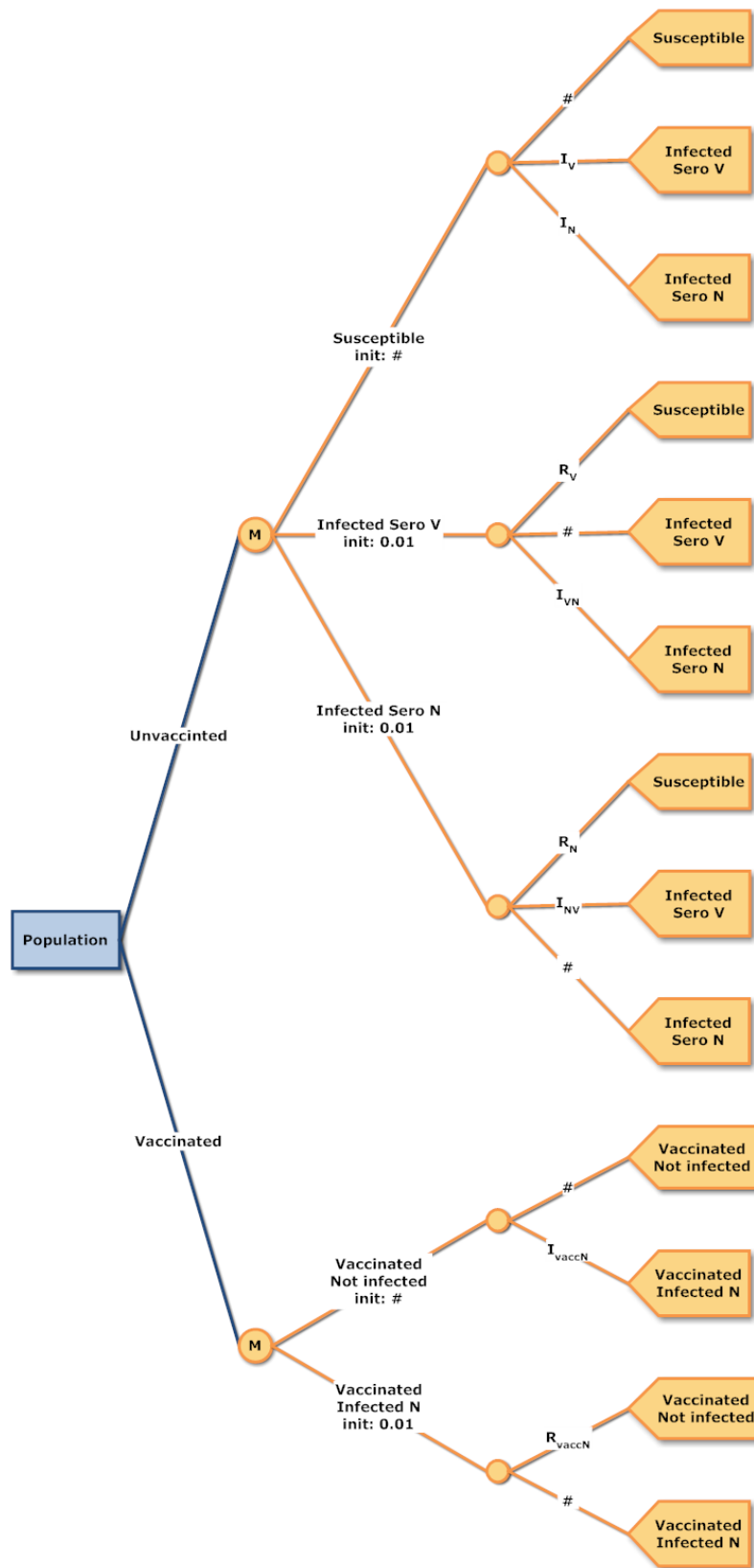
$$\tilde{M} = \begin{pmatrix} 1 - I_{vaccN} & I_{vaccN} \\ R_{vaccN} & 1 - R_{vaccN} \end{pmatrix}$$

## 8.2.2 Parameterization of the models

**Parameterization of the agent based model:** The parameterization of the agent based model is defined in chapter 8.1.3 for all comparisons with the Markov model.

**Parameterization of the Markov model:** The goal is to parameterize the Markov model so that the unvaccinated population corresponds to the base scenario of the agent based model. Therefore the following functions are defined on the unvaccinated population of the agent based model in the base scenario :  $I_V(t)$ ,  $I_N(t)$ ,  $I_{VN}(t)$ ,  $I_{NV}(t)$ ,  $R_V(t)$  and  $R_N(t)$ . The coefficients of the Markov model correspond to the quantities of these functions in the equilibrium.

## 8.2. COMPARISON OF EQUILIBRIUMS IN AN SIS EPIDEMIC



**Figure 8.2:** Tree structure of a Markov model in case of two competitive serotypes

$$\begin{aligned}
 I_V(t) &= \frac{\text{Susceptibles that get infected with serotype V in time } t}{\text{Susceptibles in time } (t-1)} \\
 I_N(t) &= \frac{\text{Susceptibles that get infected with serotype N in time } t}{\text{Susceptibles in time } (t-1)} \\
 I_{VN}(t) &= \frac{\text{Carriers with serotype V that get infected with serotype N in time } t}{\text{Carriers of serotype V in time } (t-1)} \\
 I_{NV}(t) &= \frac{\text{Carriers with serotype N that get infected with serotype V in time } t}{\text{Carriers of serotype N in time } (t-1)} \\
 R_V(t) &= \frac{\text{Carriers with serotype V that get susceptible in time } t}{\text{Carriers of serotype V in time } (t-1)} \\
 R_N(t) &= \frac{\text{Carriers with serotype N that get susceptible in time } t}{\text{Carriers of serotype N in time } (t-1)}
 \end{aligned}$$

The calculations of  $I_{vaccN}$  and  $R_{vaccN}$  require further research. The goal is that the vaccinated part of the population receive the same disadvantage by serotype N as the unvaccinated part of the population. This means that new infections and carrier with serotype N must agree with those among the unvaccinated population. This can be reached if the sum of all transition probabilities that lead to or away from a state with carriers of serotype N are the same for both population parts. This leads to the following calculations:

$$\begin{aligned}
 I_{vaccN}(t) &= \frac{\text{Susceptible people and carriers with serotype V that get infected with serotype N in time } t}{\text{Susceptible people and carriers with serotype V in time } (t-1)} \\
 R_{vaccN}(t) &= \frac{\text{Carriers of serotype N that either recover or get infected with serotype V in time } t}{\text{Carriers of serotype N in time } (t-1)}
 \end{aligned}$$

The calculations yield the following transition matrices so that the Markov model reaches the same equilibrium as the agent based model:

$$M_1 = \begin{pmatrix} 0.8709 & 0.0835 & 0.0456 \\ 0.0757 & 0.8767 & 0.0476 \\ 0.0592 & 0.0854 & 0.8554 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0.9533 & 0.0467 \\ 0.1446 & 0.8554 \end{pmatrix}$$

**Results:** Indeed, both models yield the same equilibriums. The total carrier rate for serotype V is 40.59% and the one for serotype N is 24.41%. The detailed results are explained and showed subsequently together with the test case simulations with vaccinations.



### 8.2.3 Implementation of vaccinations

**Representation of the results:** To remain comparable, it is important to define the quantities that represent the results of the simulation of both models. Due to the construction of the Markov model, it seems practicable to use the carrier rates as relevant results. They are shown in two ways: As carrier rates of Serotype V and N in the whole population and separately for the unvaccinated and vaccinated population part.

**Simulations of vaccinations:** The results of the agent based model have been discussed already in previous chapters: vaccinations cause a significant decrease of serotype V and a slight increase of serotype N. That is different for the Markov model where carrier rates of the unvaccinated and vaccinated population part remains constant for all simulations. Only the total carrier rates are changing because the parts of the populations are changing due to vaccinations.

The detailed results of the simulations with and without vaccinations are shown in figure 8.3 and in the tables 8.1 and 8.2.

### 8.2.4 Explanation of the differences and harmonization of the results

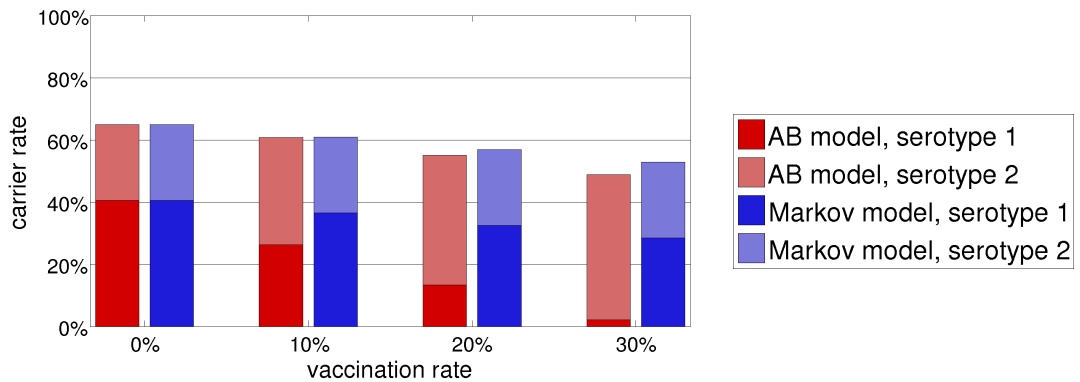
In the last step, the results of the Markov model are further processed. In this adjusted Markov model, the carrier rate of V in the unvaccinated population is reduced by the herd immunity and the carrier rates of N are increased by the serotype shift. It is important to calculate the serotype shift based on the same population as the carrier rate of serotype N. Reduction by the herd immunity means to multiply the carrier rate of V with  $(1 - H)$ , and increase of the serotype shift is to multiply the carrier rate of N with  $(1 - S)$  or with  $(1 - S_N)$  and  $(1 - S_V)$ , depending on the considered population part. This adjustment finally leads to the same results as in the agent based model.

The herd immunity and the serotype shift is shown in figure 8.4 and table 8.3 while the results of the adjusted Markov model are shown in figure 8.5 and in the referenced tables 8.1 and 8.2.

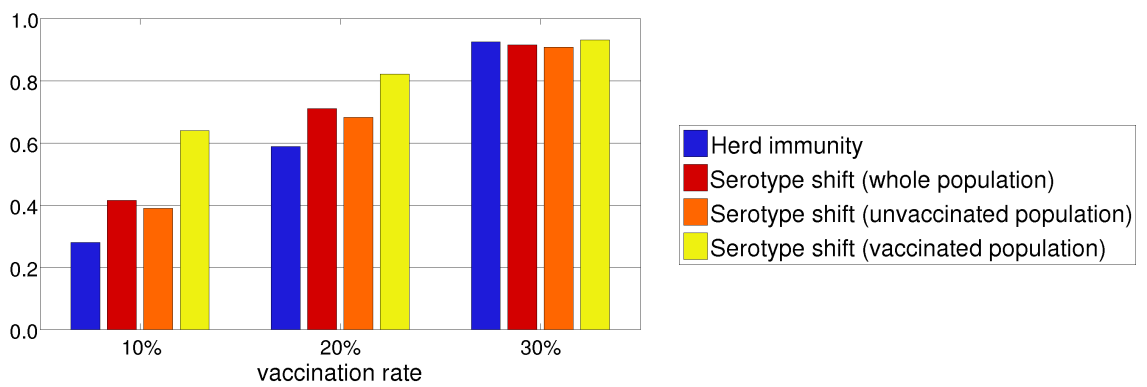
## 8.3 Comparison of the temporal progress of an SIR epidemic

### 8.3.1 Construction of the models

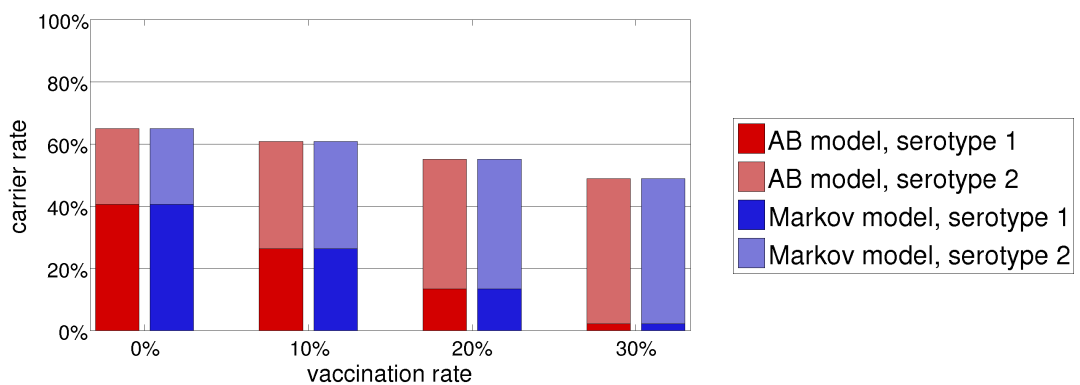
**The agent based model:** The agent based model calculates an SIR epidemic and is defined in chapter 8.1.3.



**Figure 8.3:** Carrier rates of the agent based model compared to the Markov model in the equilibrium of the SIS simulation



**Figure 8.4:** Herd immunity and serotype shift in the equilibrium of the SIS model



**Figure 8.5:** Carrier rates of the agent based model compared to the adjusted Markov model in the equilibrium of the SIS simulation

**Table 8.1:** Carrier rates in the equilibrium of the SIS simulation (separately evaluated for the unvaccinated and the vaccinated population)

Vaccination rate		Unvaccinated population, Serotype V	Unvaccinated population, serotype N	Vaccinated population, Serotype N
AB model	0%	40.59%	24.41%	—
	10%	29.23%	33.93%	40.02%
	20%	16.69%	41.07%	44.45%
	30%	3.04%	46.58%	47.14%
Markov model	0%	40.59%	24.41%	—
	10%	40.59%	24.41%	24.41%
	20%	40.59%	24.41%	24.41%
	30%	40.59%	24.41%	24.41%
Adjusted Markov model	0%	40.59%	24.41%	—
	10%	$40.59\% \cdot (1 - 0.2798)$ = 29.23%	$24.41\% \cdot (1 + 0.3901)$ = 33.93%	$24.41\% \cdot (1 + 0.6393)$ = 40.02%
	20%	$40.59\% \cdot (1 - 0.5888)$ = 16.69%	$24.41\% \cdot (1 + 0.6824)$ = 41.07%	$24.41\% \cdot (1 + 0.8210)$ = 44.45%
	30%	$40.59\% \cdot (1 - 0.9250)$ = 3.04%	$24.41\% \cdot (1 + 0.9080)$ = 46.58%	$24.41\% \cdot (1 + 0.9309)$ = 47.14%

**Table 8.2:** Carrier rates in the equilibrium of the SIS simulation (based on the total population)

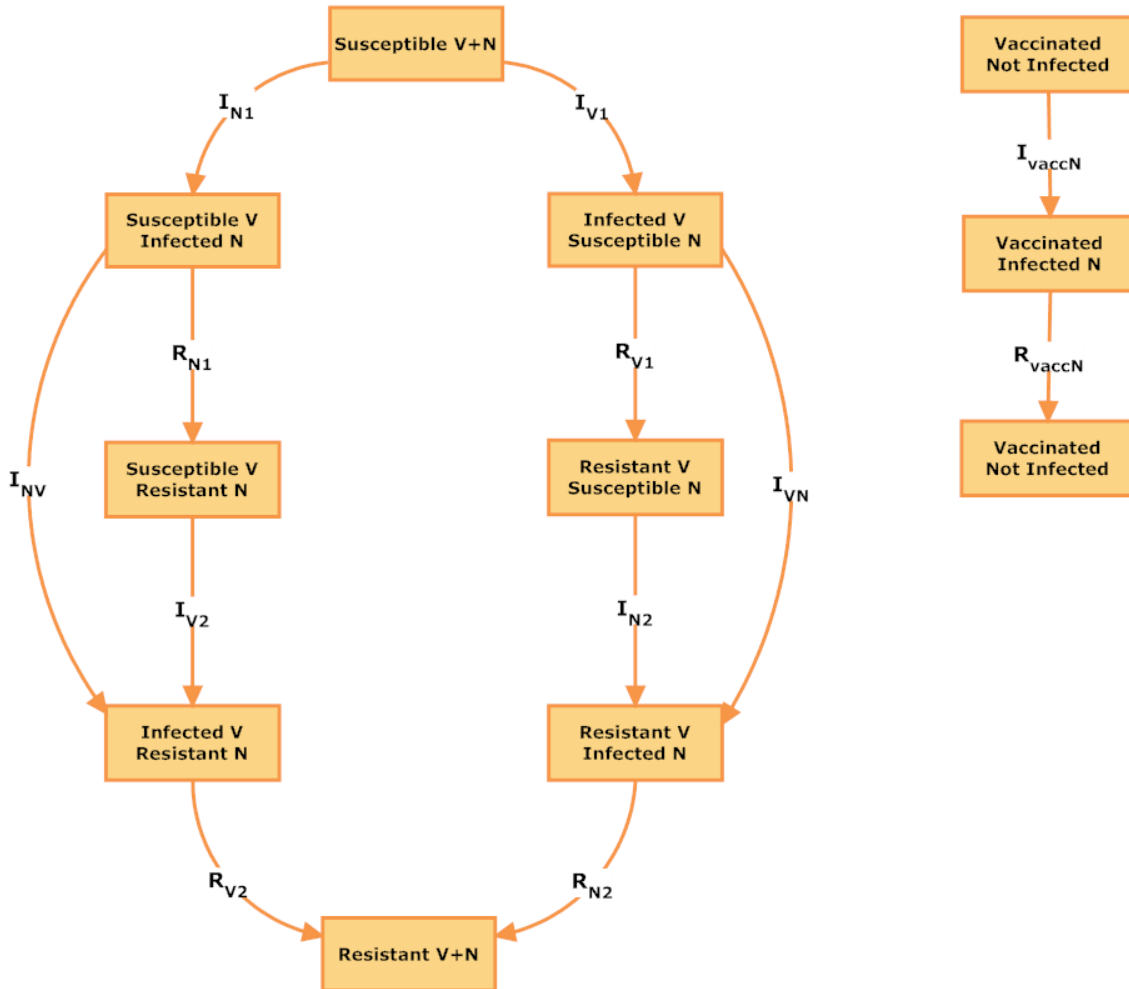
Vaccination rate		Serotype V	Serotype N
AB model	0%	$40.59\% \cdot 1 = 40.59\%$	$24.41\% \cdot 1 = 24.41\%$
	10%	$29.23\% \cdot 0.9 = 26.31\%$	$33.93\% \cdot 0.9 + 40.02\% \cdot 0.1 = 34.54\%$
	20%	$16.69\% \cdot 0.8 = 13.35\%$	$41.07\% \cdot 0.8 + 44.45\% \cdot 0.2 = 41.75\%$
	30%	$3.04\% \cdot 0.7 = 2.31\%$	$46.58\% \cdot 0.7 + 47.14\% \cdot 0.3 = 46.74\%$
Markov model	0%	$40.59\% \cdot 1 = 40.59\%$	$24.41\% \cdot 1 = 24.41\%$
	10%	$40.59\% \cdot 0.9 = 36.53\%$	$24.41\% \cdot 0.9 + 24.41\% \cdot 0.1 = 24.41\%$
	20%	$40.59\% \cdot 0.8 = 32.48\%$	$24.41\% \cdot 0.8 + 24.41\% \cdot 0.2 = 24.41\%$
	30%	$40.59\% \cdot 0.7 = 28.42\%$	$24.41\% \cdot 0.7 + 24.41\% \cdot 0.3 = 24.41\%$
Adjusted Markov model	0%	40.59%	24.41%
	10%	$36.53\% \cdot (1 - 0.2798) = 26.31\%$	$24.41\% \cdot (1 + 0.4150) = 34.54\%$
	20%	$32.48\% \cdot (1 - 0.5888) = 13.35\%$	$24.41\% \cdot (1 + 0.7101) = 41.75\%$
	30%	$28.42\% \cdot (1 - 0.9250) = 2.31\%$	$24.41\% \cdot (1 + 0.9149) = 46.74\%$

**Table 8.3:** Herd immunity and serotype shift in the equilibrium of the SIS model

Vaccination rate	Herd immunity	Serotype shift (whole population)	Serotype shift (unvaccinated population)	Serotype shift (vaccinated population)
10%	0.2798	0.4150	0.3901	0.6393
20%	0.5888	0.7101	0.6824	0.8210
30%	0.9250	0.9149	0.9080	0.9309

### 8.3. COMPARISON OF THE TEMPORAL PROGRESS OF AN SIR EPIDEMIC

**The Markov model:** The Markov model for SIR epidemics with two competitive serotypes is already very large and complex because of the relevant order of infections in this case. The state transition diagrams in figure 8.6 show the states and the transitions for unvaccinated people (on the left) and for vaccinated people (on the right). The transition probabilities are time-dependent and correspond to the notations in the state transition diagram. The tree is not shown any more due to the large size in this case. The model is defined by the state transition diagram anyway.



**Figure 8.6:** The state transition diagram of the Markov model for SIR epidemics with two competitive serotypes.

**Left:** For unvaccinated people. **Right:** For vaccinated people.

**Unvaccinated people:** Figure 8.7 shows the transition matrix and the distribution  $\underline{a} = (a_{1,t}, \dots, a_{8,t})$  of the Markov model corresponds to the 8 states:

- $s_1 =$  susceptible V+N
- $s_2 =$  susceptible V & infected N
- $s_3 =$  susceptible V & resistant N
- $s_4 =$  infected V & resistant N

$s_5$  = infected V & susceptible N  
 $s_6$  = resistant V & susceptible N  
 $s_7$  = resistant V & infected N  
 $s_8$  = resistant V+N

**Vaccinated people:** The distribution  $\tilde{a} = (\tilde{a}_{1,t}, \tilde{a}_{2,t}, \tilde{a}_{3,t})$  of the Markov model contains three elements according to the state set  $\tilde{S} \in \tilde{S}$  and figure 8.8 shows the transition matrix.

$\tilde{s}_1$  = vaccinated & susceptible N  
 $\tilde{s}_2$  = vaccinated & infected N  
 $\tilde{s}_3$  = vaccinated & resistant N

### 8.3.2 Parameterization of the models

**Parameterization of the agent based model:** The parameterization of the agent based model is defined in chapter 8.1.3 for all comparisons with the Markov model.

**Parameterization of the Markov model:** The Markov model has to be parameterized so that its temporal progress coincides with the agent based model. This task requires time-dependent transition matrices. The definitions of  $M_1(t)$  and  $M_2(t)$  in the previous chapter (figures 8.7 and 8.8) are able to do this.

The transition probabilities  $I_{N1}(t)$ ,  $R_{N1}(t)$ ,  $I_{NV}(t)$ ,  $I_{V2}(t)$ ,  $R_{V2}(t)$ ,  $I_{V1}(t)$ ,  $R_{V1}(t)$ ,  $I_{VN}(t)$ ,  $I_{N2}(t)$  and  $R_{N2}(t)$  of the unvaccinated population part are calculated for each time step in the agent based model. The transition probability in time  $t$  calculates as the part of persons in a certain state in time  $t$  who change into another state in the subsequent time ( $t + 1$ ). The transition probabilities  $I_{vaccN}(t)$  and  $R_{vaccN}(t)$  for the vaccinated population part are set in a way that their infection rates and recovery rates with serotype N are the same as in the unvaccinated population.

**Representation of the results:** The results are represented by the temporal progression of the carrier rates in the whole population. The carrier rates with serotype N in the vaccinated and unvaccinated population part is very similar, hence they are not shown for the sake of clarity in the figures.

**Results:** Figure 8.9 shows that the epidemic curves indeed coincide in the base case without vaccinations. The results of the Markov models are shown as ‘x’ markers while those of the agent based model are plotted as solid lines.

$$M = \begin{pmatrix} 1 - I_{N1}(t) - I_{V1}(t) & I_{N1}(t) & 0 & 0 & I_{V1}(t) & 0 & 0 & 0 \\ 0 & 1 - R_{N1}(t) - I_{NV}(t) & R_{N1}(t) & I_{NV}(t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - I_{V2}(t) & I_{V2}(t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - R_{V2}(t) & 0 & 0 & 0 & R_{V2}(t) \\ 0 & 0 & 0 & 0 & 1 - R_{V1}(t) - I_{VN}(t) & R_{V1}(t) & I_{VN}(t) & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 - I_{N2}(t) & I_{N2}(t) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 - R_{N2}(t) & R_{N2}(t) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

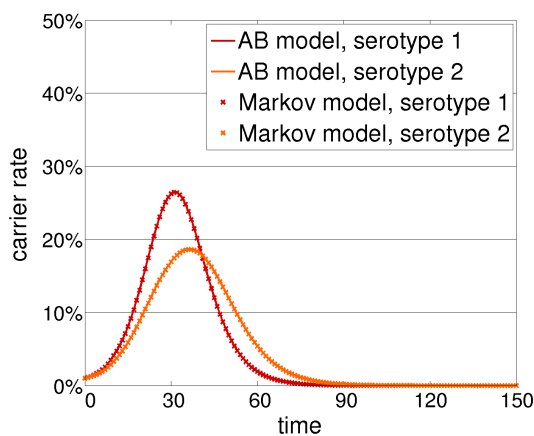
**Figure 8.7:** Transition matrix of the Markov model for SIR epidemics with two competitive serotypes for unvaccinated people

$$\tilde{M} = \begin{pmatrix} 1 - I_{vaccN}(t) & I_{vaccN}(t) & 0 \\ 0 & 1 - R_{vaccN}(t) & R_{vaccN}(t) \\ 0 & 0 & 1 \end{pmatrix}$$

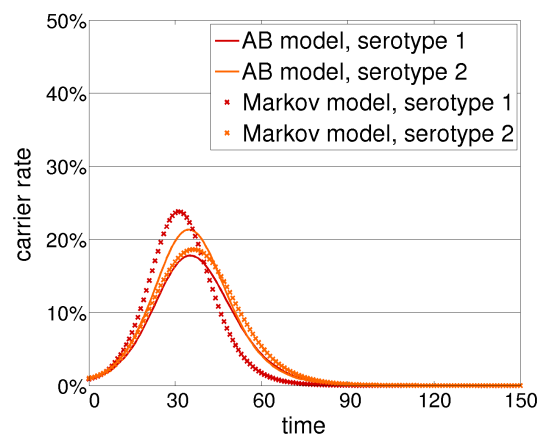
**Figure 8.8:** Transition matrix of the Markov model for SIR epidemics with two competitive serotypes for vaccinated people

### 8.3.3 Implementation of vaccinations

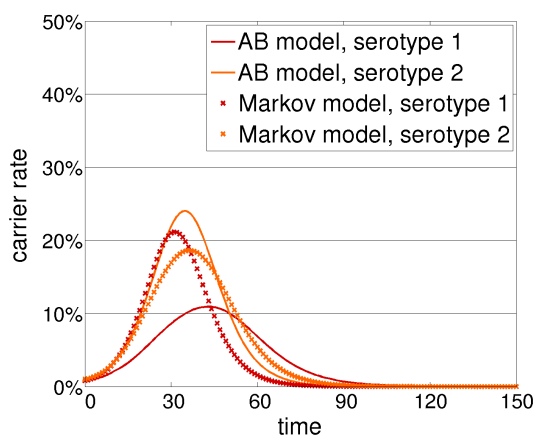
The vaccinations in the models are performed as explained in the previous chapters. The carrier rates are shown in figures 8.10, 8.11 and 8.12 for 10%, 20% and 30% vaccinations. It turns out that serotype V (red) is too strong in the Markov model while serotype N (orange) is too weak compared to the agent based model in all three cases. This behavior can easily be explained by herd immunity which affects serotype V, and serotype shift which affects serotype N.



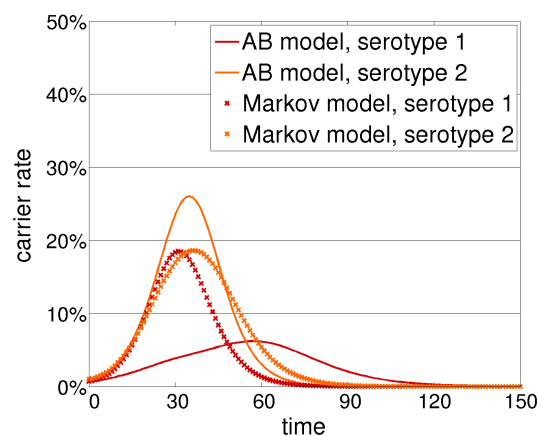
**Figure 8.9:** The carrier rates of the agent based and the Markov model for SIR epidemics in the base case without vaccinations



**Figure 8.10:** The carrier rates of the agent based and the Markov model for SIR epidemics in the test case with 10% vaccinations



**Figure 8.11:** The carrier rates of the agent based and the Markov model for SIR epidemics in the test case with 20% vaccinations

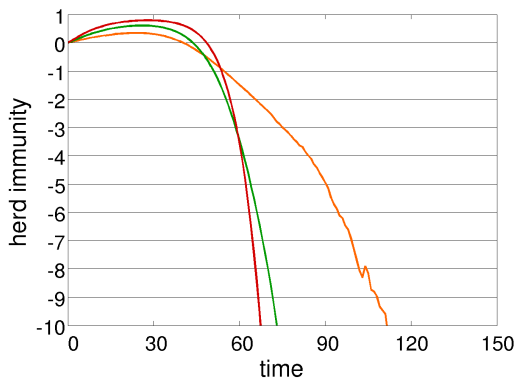


**Figure 8.12:** The carrier rates of the agent based and the Markov model for SIR epidemics in the test case with 30% vaccinations

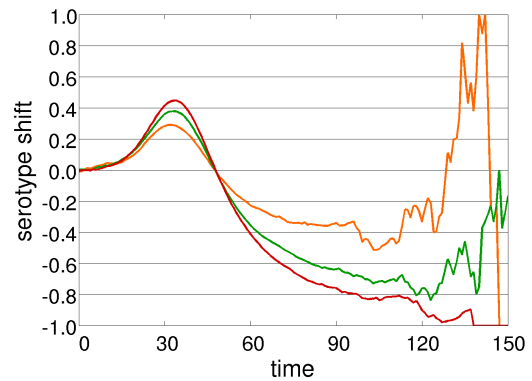


### 8.3.4 Explanation of the differences and harmonization of the results

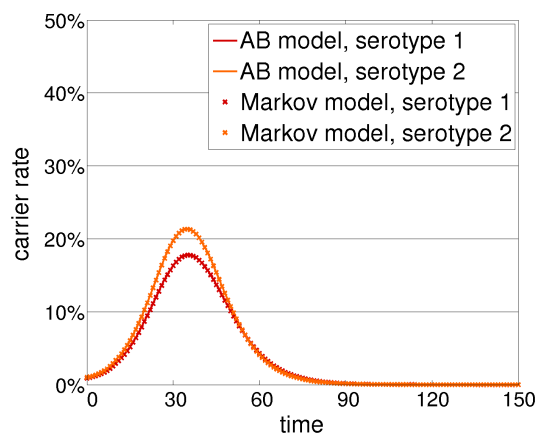
The herd immunity and the serotype shift are applied on the results of the Markov model as a correction. This means that the carrier rate with serotype V is increased by the herd immunity and the one with serotype N is decreased by the serotype shift. Both epidemic effects are calculated upon the current carrier rate in the agent based model. They are plotted in the figures 8.13 and 8.14. The figures 8.15, 8.16 and 8.17 show that the corrected Markov model again yields the same results as the agent based model.



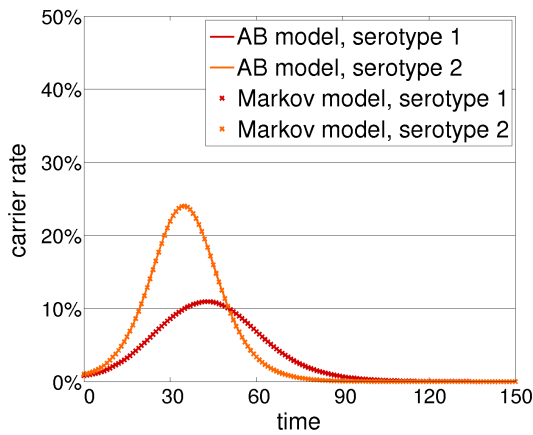
**Figure 8.13:** Herd immunity in the SIR model



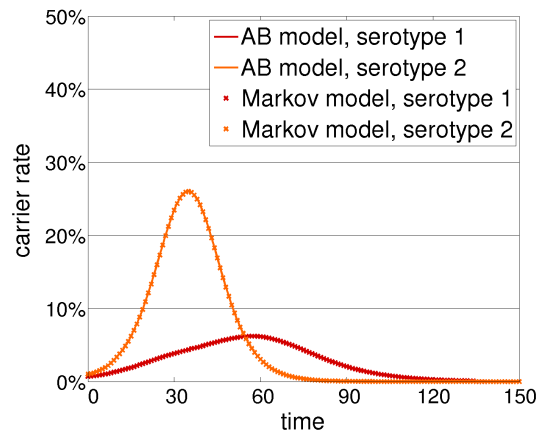
**Figure 8.14:** Serotype shift in the SIR model



**Figure 8.15:** The carrier rates of the agent based and the corrected Markov model for SIR epidemics in the test case with 10% vaccinations



**Figure 8.16:** The carrier rates of the agent based and the corrected Markov model for SIR epidemics in the test case with 20% vaccinations



**Figure 8.17:** The carrier rates of the agent based and the corrected Markov model for SIR epidemics in the test case with 30% vaccinations

## 8.4 Comparison of a whole SIR epidemic

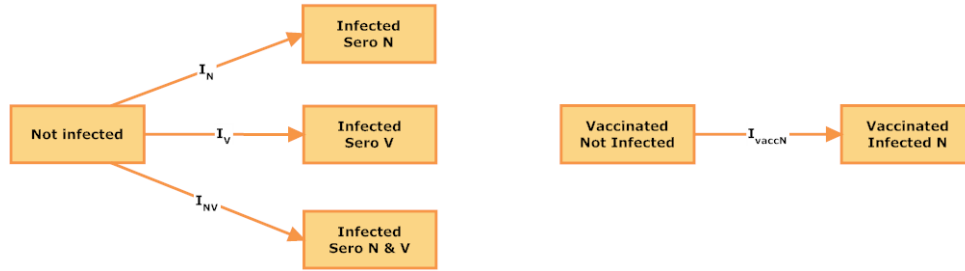
**Remark:** SIR epidemics are often not simulated over time but are instead considered as one event that represents the whole season from appearance of the pathogen until extinction. The results are represented by the total number of infections within the whole population, and separately within the vaccinated and unvaccinated population part.

### 8.4.1 Construction of the models

**The agent based model:** The agent based model is the same as used in the previous chapters. It is defined in chapter 8.1.3. The results of this model are not time-dependent. Instead, they are the sum of infections of the whole simulation.

**The Markov model:** The Markov model is very simple. The whole epidemic is summed up in a single event, hence the Markov model needs to simulate only one time step. The states represent what infections happened to the persons during the whole season. There are four possibilities: not infected, infected with serotype V, infected with serotype N, infected with both serotypes. This leads to the structure of the model which is shown in figure 8.18 by the state transition diagram.

It is important to mention that not the states but the numbers of infections are the results of the Markov model. The infections with serotype N in the unvaccinated population are the sum of the states “Sero N” and “Sero N & V” and in the vaccinated population “Vaccinated Infected N”. The infections with serotype V are calculated analogously.



**Figure 8.18:** State transition diagram of the Markov Model of a whole SIR season. **Left:** For unvaccinated people. **Right:** For vaccinated people.

### 8.4.2 Parameterization of the models

**Parameterization of the agent based model:** The parameterization of the agent based model is defined in chapter 8.1.3 for all comparisons with the Markov model.

**Parameterization of the Markov model:** For the initial configuration, 98% of the people are considered to be “not infected” and 1% each with serotype V and N. To reproduce the results of the agent based model one has to count the number of people who get infected during the whole season. Based on these numbers, the transition probabilities in the Markov model are calculated. The vaccinated population part should have the same disadvantage by serotype N as the unvaccinated part, hence  $I_{vaccN} = I_N + I_{NV}$  is defined. Table 8.4 shows this calculations.

**Table 8.4:** Parameterization of the Markov model of a whole SIR season by the agent based model

	Infections in the agent based model	Transition probabilities of the Markov model
unvaccinated population	Serotype V: 19.39% Serotype N: 13.96% Serotype N & V: 59.76%	$I_V = 0.1939$ $I_N = 0.1396$ $I_{NV} = 0.5976$
vaccinated population	- -	$I_{vaccN} = 0.1396 + 0.5976 = 0.7372$

**Results:** It turns out that both serotypes can infect more than 70% of the population each during a whole season, with serotype V causing slightly more infections. The Markov model yields the same results with the calculated parameters. The detailed numbers are shown in the tables 8.5 and 8.6 in the following chapter together with the results of the vaccinations.

### 8.4.3 Implementation of vaccinations

Implementation of vaccinations in the agent based model shows the well known behavior of SIR epidemics: a strong decrease of serotype V and a slight increase of serotype N. The Markov model calculates too many infections with serotype V and too few with serotype N. The results are shown in the tables 8.5 and 8.6.

**Table 8.5:** Total infections of an SIR epidemic (based on the unvaccinated and vaccinated part of the population)

Vaccination rate		Unvaccinated population, Serotype V	Unvaccinated population, Serotype N	Vaccinated population, Serotype N
AB model	0%	79.15%	73.72%	—
	10%	73.32%	74.44%	74.77%
	20%	65.95%	75.77%	75.99%
	30%	56.00%	77.26%	77.41%
Markov model	0%	79.15%	73.72%	—
	10%	79.15%	73.72%	73.72%
	20%	79.15%	73.72%	73.72%
	30%	79.15%	73.72%	73.72%
Adjusted Markov model	0%	79.15%	73.72%	—
	10%	$79.15\% \cdot (1 - 0.0737)$ = 73.32%	$73.72\% \cdot (1 + 0.0098)$ = 74.44%	$73.72\% \cdot (1 + 0.0143)$ = 74.77%
	20%	$79.15\% \cdot (1 - 0.1668)$ = 65.95%	$73.72\% \cdot (1 + 0.0278)$ = 75.77%	$73.72\% \cdot (1 + 0.0308)$ = 75.99%
	30%	$79.15\% \cdot (1 - 0.2926)$ = 56.00%	$73.72\% \cdot (1 + 0.0481)$ = 77.26%	$73.72\% \cdot (1 + 0.0500)$ = 77.41%

### 8.4.4 Explanation of the differences and harmonization of the results

The herd immunity and the serotype shift are calculated in the agent based model upon the sum of infections of the whole simulation. They are shown in table 8.7. In both cases, for the vaccinated and unvaccinated population as well as for the whole population, the results of the adjusted Markov model are the same as of the agent based model (see the results tables 8.5 und 8.6). Hence, the concept also works for longer time periods.

**Table 8.6:** Total infections of an SIR epidemic (based on the whole population)

Vaccination rate		Serotype V	Serotype N
AB model	0%	$79.15\% \cdot 1 = 79.15\%$	$73.72\% \cdot 1 = 73.72\%$
	10%	$73.32\% \cdot 0.90 = 65.99\%$	$74.44\% \cdot 0.90 + 74.77\% \cdot 0.10 = 74.48\%$
	20%	$65.95\% \cdot 0.80 = 52.76\%$	$75.77\% \cdot 0.80 + 75.99\% \cdot 0.20 = 75.81\%$
	30%	$56.00\% \cdot 0.70 = 39.20\%$	$77.26\% \cdot 0.70 + 77.41\% \cdot 0.30 = 77.31\%$
Markov model	0%	79.15%	73.72%
	10%	$79.15\% \cdot 0.90 = 71.24\%$	$73.72\% \cdot 0.90 + 73.72\% \cdot 0.10 = 73.72\%$
	20%	$79.15\% \cdot 0.80 = 63.32\%$	$73.72\% \cdot 0.80 + 73.72\% \cdot 0.20 = 73.72\%$
	30%	$79.15\% \cdot 0.70 = 55.41\%$	$73.72\% \cdot 0.70 + 73.72\% \cdot 0.30 = 73.72\%$
Adjusted Markov model	0%	79.15%	73.72%
	10%	$71.24\% \cdot (1 - 0.0737) = 65.99\%$	$73.72\% \cdot (1 + 0.0103) = 74.48\%$
	20%	$63.32\% \cdot (1 - 0.1668) = 52.76\%$	$73.72\% \cdot (1 + 0.0284) = 75.81\%$
	30%	$55.41\% \cdot (1 - 0.2926) = 39.20\%$	$73.72\% \cdot (1 + 0.0487) = 77.31\%$

**Table 8.7:** Herd immunity and serotype shift of a whole SIR epidemic

Vaccination rate	Herd immunity	Serotype shift (whole population)	Serotype shift (unvaccinated population)	Serotype shift (vaccinated population)
10%	0.0737	0.0103	0.0098	0.0143
20%	0.1668	0.0284	0.0278	0.0308
30%	0.2926	0.0487	0.0481	0.0500

## 8.5 Analytic considerations

In all three examples the differences between the agent based and the Markov model were exactly given by herd immunity and serotype shift. It is even possible to show these relations in general.

**Assumptions:** The basis are an agent based and a Markov model. The results of both models must be represented by time-dependent functions:

- $T_{\{Model\},\{Scenario\},N,V}(t)$ :  
The disadvantage of the unvaccinated population part by serotype V
- $T_{\{Model\},\{Scenario\},N,N}(t)$ :  
The disadvantage of the unvaccinated population part by serotype N
- $T_{\{Model\},\{Scenario\},V,N}(t)$ :  
The disadvantage of the vaccinated population part by serotype N

The parameterization of the agent based model is given for the base scenario. Then, the Markov model must be parameterized so that its results equals the agent based model. If there are no vaccinated persons in the agent based model, then the parameterization of the Markov model has to be chosen so that  $T_{Markov,\{Scenario\},V,N}(t) = T_{Markov,\{Scenario\},N,N}(t)$ .

**Vaccinations:** The test scenario yields the new results  $T_{AB,test,N,V}(t)$ ,  $T_{AB,test,N,N}(t)$  and  $T_{AB,test,V,N}(t)$  of the agent based model. These quantities allow to calculate herd immunity  $H(t)$  and the serotype shift  $S_N(t)$  of the unvaccinated population and  $S_V(t)$  of the vaccinated population by definition:

$$H(t) = \frac{T_{AB,base,N,V}(t) - T_{AB,test,N,V}(t)}{T_{AB,base,N,V}(t)}$$

$$S_N(t) = \frac{T_{AB,test,N,N}(t) - T_{AB,base,N,N}(t)}{T_{AB,base,N,N}(t)}$$

$$S_V(t) = \frac{T_{AB,test,V,N}(t) - T_{AB,base,N,N}(t)}{T_{AB,base,N,N}(t)}$$

Vaccinations in the Markov model only change the distribution of the cohorts, but within cohorts the relative disadvantage remains unchanged. This means that:

$$T_{Markov,base,N,V}(t) = T_{Markov,test,N,V}(t)$$

$$T_{Markov,base,N,N}(t) = T_{Markov,test,N,N}(t)$$

$$T_{Markov,base,V,N}(t) = T_{Markov,test,V,N}(t)$$

**Analyzation of the results:** Under these assumptions, it is easy to how how the results of both models are related.

**Relation 1: Decrease of the disadvantage by serotype V of the unvaccinated population in the Markov model by the herd immunity yields the corresponding result of the agent based model.**

$$\begin{aligned}
T_{Markov,test,N,V}(t) \cdot (1 - H(t)) &= \\
&= T_{Markov,base,N,V}(t) \cdot \left(1 - \frac{T_{AB,base,N,V}(t) - T_{AB,test,N,V}(t)}{T_{AB,base,N,V}(t)}\right) = \\
&= T_{AB,base,N,V}(t) \cdot \left(\frac{T_{AB,test,N,V}(t)}{T_{AB,base,N,V}(t)}\right) = \\
&= T_{AB,test,N,V}(t)
\end{aligned}$$

**Relation 2: Increase of the disadvantage by serotype N of the unvaccinated population in the Markov model by the serotype shift yields the corresponding result of the agent based model.**

$$\begin{aligned}
T_{Markov,test,N,N}(t) \cdot (1 + S_N(t)) &= \\
&= T_{Markov,base,N,N}(t) \cdot \left(1 + \frac{T_{AB,test,N,N}(t) - T_{AB,base,N,N}(t)}{T_{AB,base,N,N}(t)}\right) = \\
&= T_{AB,base,N,N}(t) \cdot \left(\frac{T_{AB,test,N,N}(t)}{T_{AB,base,N,N}(t)}\right) = \\
&= T_{AB,test,N,N}(t)
\end{aligned}$$

**Relation 3: Increase of the disadvantage by serotype N of the vaccinated population in the Markov model by the serotype shift yields the corresponding result of the agent based model.** The corresponding calculation works analogously to the one in relation 2.

**Remark:** The disadvantage does not necessarily need to be considered separately within the vaccinated and the unvaccinated population. It is possible to show in the same way that the disadvantage of the whole population or of other population parts can also be corrected by the herd immunity and the serotype shift. In this case, both epidemic effects  $H(t)$  and  $S(t)$  needs to be calculated on the respective population. Hence, the relationship between Markov models and agent based models for epidemics is generally shown.

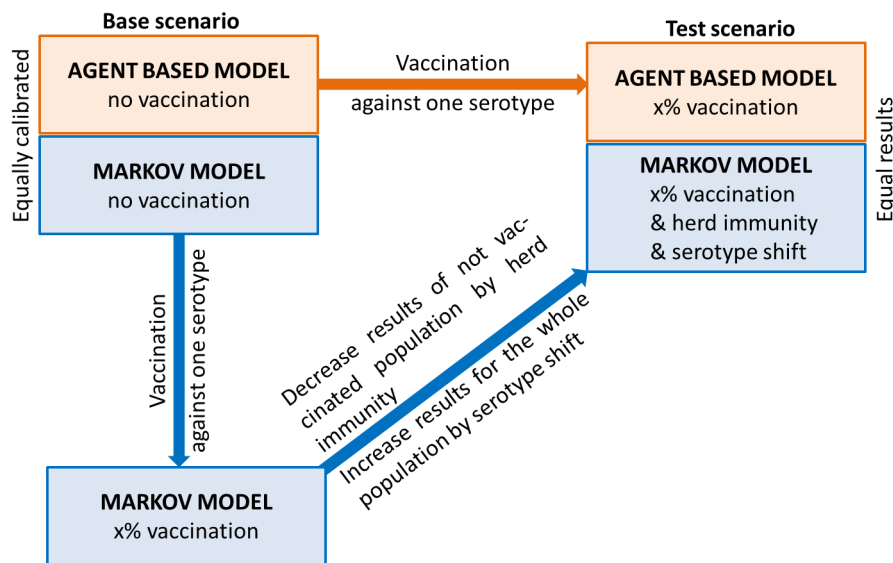
## 8.6 Conclusions

A comparison between Markov models and agent based models is also possible for epidemics with two competitive serotypes. For this purpose, the simulation results have to be differentiated by the serotypes. Furthermore, the results of both models must be structurally comparable. The herd immunity in a scenario with vaccinations represents

CHAPTER 8. COMPARISON OF MARKOV AND AGENT BASED MODELS FOR EPIDEMIC SIMULATIONS WITH TWO COMPETITIVE SEROTYPES

the benefit of unvaccinated persons concerning this serotype, while serotype shift represents an increased disadvantage of the whole population concerning the other serotype. These quantities can always be calculated in the agent based model. The Markov model behaves the same, independent of the vaccination rate. Hence, it differs from the agent based model exactly by the herd immunity and the serotype shift.

This circumstance is shown in three examples and is also analytically proven. Hence, the definition of herd immunity and serotype shift in chapter 7 is suitable to represent, interpret and use the emerging behavior of a contact based epidemic model. Figure 8.19 shows this situation.



**Figure 8.19:** Relationship of vaccinations between agent based models and Markov models with two serotypes.



# Chapter 9

## Summary and Conclusions

### 9.1 Summary of insights

This work deals with the spread of infectious diseases. In contrast to other studies in this area, it does not examine real populations or patients in clinical trials. Instead, it uses models as tools to answer the questions. Advantages of models are that they are cheaper, allow wide examinations of details that are not always possible in real populations and that they can be performed in an endless number of variations. The disadvantage is that models only represent a simplified image of reality but cannot reflect reality completely. Critical views on the model and extensive validation allow to keep the deviation between model and reality small.

There are numerous possibilities to simulate infectious diseases and associated epidemics. The chapters two and three contain short introductions into some important techniques with a special focus on agent based model. The agent based models allow, due of their close-to-reality-structure, simulation of spread of diseases as a consequence of transmissions from person to person. This way of modeling often leads to several interesting effects. Agent based models are in contrast to commonly used Markov models which are in fact an extrapolation and rearrangement of known data. Effects that occur in agent based models can theoretically be considered in Markov models, but they must be known completely and quantitatively.

The main goal of this work is to define, to examine and interpret two common effects in the area of epidemiology: **herd immunity** and **serotype shift**.

Herd immunity is already a household name since the 1920s, but its occurrence has always been very disputed. Literature research reveals that there is not even a consensus about what herd immunity actually is. Still, it is easy to explain the background of the phenomenon “herd immunity”: If a part of a population gets vaccinated against an infectious disease then not vaccinated persons experience a benefit. These unvaccinated persons meet potentially less infectious persons (because some of them are vaccinated) so they have a lower probability to get infected. This effect can even lead to extinction of a

disease even though only a part of the population is vaccinated. These considerations lead to the main definition of this study: Herd immunity is the benefit for not vaccinated people by a vaccination strategy as an intervention in the system. The benefit itself is defined as the reduction of disadvantage caused by the disease in comparison of two scenarios. This leads to the problem that the disadvantage of a disease can be determined in many ways. New infections, carrier rates, symptomatic cases or economic outcomes are only a few values that are potentially usable. Analytic examinations and parameter studies show that different calculations turn out to be appropriate in different situations to express herd immunity in a useful way. The right calculation method should be chosen individually for every single case. It highly depends on what characterizes the disadvantage of a particular epidemic best and which values are measurable. The temporal component has also an effect. It is possible to calculate the disadvantage in a single point in time or accumulated over a period of time or since simulation start. Any temporal consideration seems suitable for steady epidemics. But, interventions on periodic or time limited epidemics can cause temporal offsets or distortions. Useful interpretations of herd immunity in such cases can only be made with an accumulated calculation of the disadvantage. Anyway, the herd immunity is a time-dependent function which represents the benefit of persons that are not directly affected by an intervention. The examples in this study always use a homogenous population which is subject to a vaccination strategy against the pathogen of a disease. Appropriate calculation methods show that the benefit of the unvaccinated part of the population usually grows with a higher vaccination rate. Additionally, it shows that the herd immunity has a nonlinear behavior and does not only depend on the vaccination rate but also on the general configuration of the whole model. Special attention is paid to situations where a disease cannot spread any more and becomes extinct. Then the herd immunity reaches a maximum so these situations are called **complete herd immunity**.

The previous discussed aspects about spread of diseases are not always sufficient. Pathogens can mutate and then different strains, so-called serotypes, occur. Serotypes are often in a competition. This means that infected persons cannot be infected additionally with another serotype. However, under certain assumptions serotypes might be able to “kick” another serotype “out” of an infected person and colonize him or her themselves. The individuals of a population are some kind of the living space for pathogens which is divided and embattled among the serotypes. Vaccines against such diseases often fight only one or some serotypes due to their biochemical production procedure. Vaccinations of people in a population cause a completely new situation which results in a new distribution of the living space. The idea of **serotype shift** is to show and explain possible effects of changing occurrence of unaffected serotypes. Hereafter, we consider two competitive serotypes causing an epidemic. Again two scenarios are compared; serotype shift is defined as the increased disadvantage by not covered serotypes in the vaccine that affects the whole population. It is useful to calculate this effect separately within the vaccinated and unvaccinated population. Based on the previous knowledge on single-serotype epidemics, it is practicable to define **herd immunity** in such cases as the decrease of the disadvantage caused by vaccine-covered serotypes in the unvaccinated population. The **general herd immunity** considers the whole effect of a vaccination strategy on the unvaccinated population, disregarding the causing serotypes. Examinations show that the general herd

immunity is just a summarized effect composed of the herd immunity and the serotype shift on the unvaccinated population. **Complete herd immunity** refers to the situation when the vaccine-covered serotypes cannot spread any more and become extinct. This definition is consistent with the single-serotype-case because it is also characterized by the highest possible value of herd immunity. The calculation of the disadvantage is done in the same way using an appropriate value and temporal component. Parameter studies using the agent based model simulate a homogenous population, two competitive serotypes and a vaccination that fights only one serotype. All results clearly show prevalence of both herd immunity and serotype shift. This means that unvaccinated persons benefit from the vaccination due to reduction of one serotype, but the whole population suffers from an increasing disadvantage caused by the other serotype. All in all the vaccination strategies still have a positive impact, but it is often reduced by occurring serotype shift. Both effects are stronger for higher vaccination rates, but also depend on other configurations of the model. These dependencies are obviously nonlinear.

The calculation methods of herd immunity and serotype shift are so generally defined that they are also applicable for inhomogeneous populations and for interventions other than vaccinations. The only presumption is that a distinction between affected and not affected people by the intervention can be made. In the case of two or more serotypes, a distinction of the disadvantage caused by the single serotypes is also required. Consequently, herd immunity and serotype shift can potentially be calculated and analyzed in various models but also in real populations and studies.

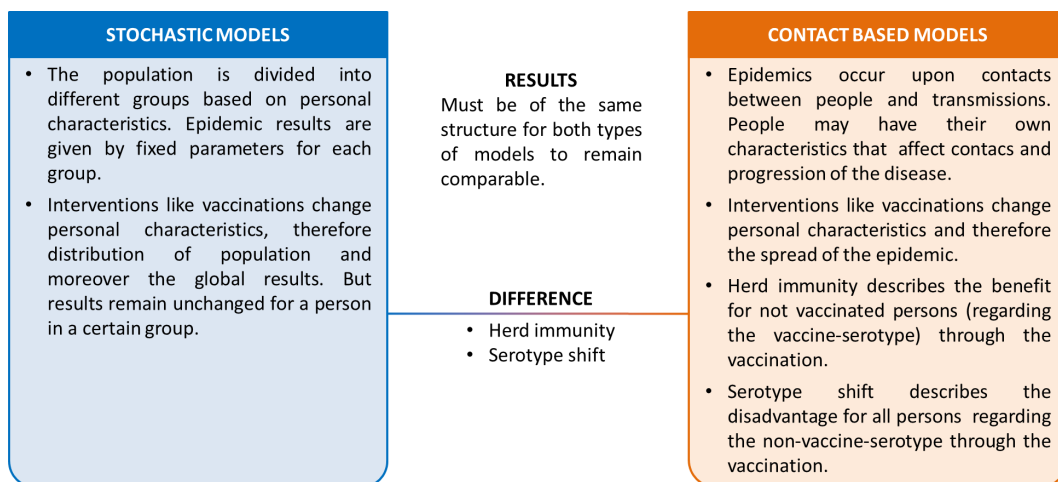
Another examination in this work is a comparison of the agent based model and a Markov model. In not less than six studies for both one and two serotypes, main characteristics of the epidemics were compared: the levels in the equilibrium of an SIS model, the time dependent spread of an SIR model and the result of a whole SIR epidemic. The first step of the analysis showed that Markov models can always produce the same results as agent based models with an appropriate structure and parameterization. Actually, the procedure is a calibration of the Markov model by choosing the right transition probabilities. Even though this method is always possible, it does not explain anything. For this reason, the second step goes beyond this. It starts with a scenario of an epidemic without vaccination. Both the agent based and the Markov models are parameterized to yield the same results. But, implementation of a vaccination strategies leads to strongly divergent results. Calculations in the single-serotype case show that this divergence exactly matches the previously defined herd immunity. This means that decrease of infections within the unvaccinated population by the herd immunity in the Markov model yield the result of the agent based model. This relation can be extended on epidemics of two competitive serotypes: The difference of the results corresponds exactly to the herd immunity and the serotype shift. This means that decrease of unvaccinated infected people with the vaccine-serotype by the herd immunity and increase of infected people with the other serotype by serotype shift in the Markov model yield the results of the agent based model. This calculations even allow correct simulation of situations with complete herd immunity with a Markov model, which means that the disease becomes extinct due to vaccinations. With support of this research it is easy to show all these relations even analytically.

## 9.2 Conclusions

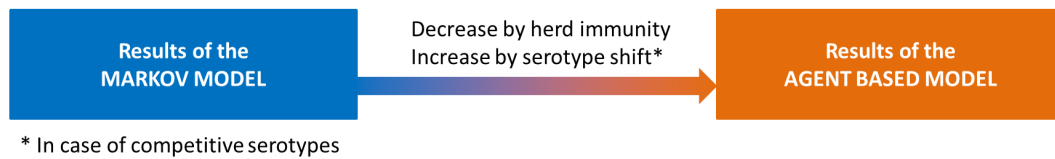
It is obvious that models that simulate epidemics based on contacts and transmissions from person to person, behave highly nonlinear. Markov models, in contrast, are linear dynamical systems, hence they yield different results for the same problem.

Herd immunity and serotype shift can be used directly to explain and show the qualitative and the quantitative benefit of unvaccinated people by a vaccination strategy, as well as the disadvantage by other serotypes that reach a higher prevalence after vaccinations. On top of that, it becomes apparent that herd immunity and serotype shift, if competitive serotypes exist, describe the whole differences between stochastic calculations and transmission based models for epidemics. Therefore, these quantities allow a complete description of all nonlinear effects on the spread of epidemics caused by contacts and transmissions. Furthermore, the research shows that herd immunity and serotype shift are not constant values, instead they are time-dependent functions that depend on the vaccination rate and on other aspects of the system configuration in a nonlinear way. Even though their origin and explanation is clear, they are often not known and cannot be measured in real world situations. Thus, it is possible to draw conclusions about usage of Markov models for prediction of disease interventions in chapter 9.4. The figures 9.1 and 9.2 outline a comparison of agent based models and Markov models for epidemic simulation.

The used definition and interpretations of herd immunity and serotype shift in this study absolutely make sense because firstly, they can explain epidemic effects that occur in populations and secondly, they lead to a deeper understanding of epidemiologic processes and their representation in models.



**Figure 9.1:** Stochastic and contact based epidemic models in comparison



**Figure 9.2:** Stochastic and contact based epidemic models in comparison

## 9.3 Further considerations

The insights lead to some general considerations on modeling of infectious diseases. Herd immunity and serotype shift are computable very well upon comparison of different situations in an agent based model. Things are very different in real systems. There, herd immunity can hardly ever be captured because testing of interventions in a closed population is usually not possible and so there is no test scenario available. Direct usage of clinical data in stochastic models is somehow unrewarding due to the shown differences to agent based models. Instead, only modeling of most realistic structures might provide satisfying predictions of effects of interventions.

The represented type of contacts in the model depend on the epidemic because they should be able to cover all transmissions of the actual disease. Plausible contact systems are for example all interpersonal meetings with a maximum distance for airborne diseases, direct or indirect physical contacts for smear infections or sexual contacts for sexually transmitted diseases.

Many ways can lead to an appropriate complexity of a modeled contact system and other epidemic influences. Increasing levels of details can be reached among other things by modeling of ...

- ... additional personal attributes like age or “general personal health”.
- ... different places where contacts happen.
- ... duration of contacts.
- ... intensity of contacts.
- ... individual contact numbers.
- ... individual transmission probabilities.

Such additional model details can be realized in any desired level of details, and they can be combined. For example, people could have age dependent contact numbers with persons of different age classes. Transmission probabilities could depend on the individual immune system of a person or of the intensity or the duration of the contact. Contact numbers and intensity could be dependent on the place.

The opportunities for higher complexity seem to be endless. In this context, an important question arises: How detailed should a model be, to be considered as an acceptable image

of reality? The main principle in modeling and simulation is a good hint: “A model should be as simple as possible but as complex as required.” A general recommendation is definitely not possible. Instead, the appropriate level of details should be evaluated by approved methods of validation for a concrete problem. [40] [48]

## 9.4 Recommendation for usage of Markov models

Markov models are widely used to evaluate aspects and changes in the field of health care. The model comparison in this study allows to determine in which situations can the usage of Markov models for epidemic simulations generally produce valid results. The problems of Markov models are, on the one hand, that they are not able to construct epidemics structurally based. Instead, they require exact, time-dependent parameterization. On the other hand, they cannot correctly simulate changes in the epidemic caused by interventions. Therefore, usage of Markov models for infectious diseases is only recommendable if all three of the following statements are true:

- The intervention does not influence the epidemic spread.
- The modeled timeframe contains either an equilibrium or a periodic epidemic propagation (for example caused by seasonality).
- If the epidemic is not in an equilibrium, the duration of a time step in the Markov model must be a full period or a multiple of a period.

For example, Markov models might be used to evaluate vaccination of a very few people. For seasonal epidemics like the influenza, it is important that the time step always cover a complete season.

Theoretically, Markov models could also be used if detailed data of herd immunity and of serotype shift are available for every simulated intervention. Generation of such data in a clinical study requires a sufficiently large, closed population which seems almost impossible because it would correspond with a “real” structural reproduction of the agent based model. Hence, usage of Markov models is actually not recommendable for simulating widespread vaccination strategies and during ongoing adaption phases.

# Bibliography

- [1] Die weitere Entwicklung ist nicht beantwortbar. - Wird die Neue Influenza über- oder unterschätzt? *Ärzte Woche* 48 /2009, Springer-Verlag GmbH, 11 2009. URL <http://www.springermedizin.at/artikel/14179-die-weitere-entwicklung-ist-nicht-beantwortet>.
- [2] W. Atkinson, C. Wolfe, and J. Hamborsky. *Epidemiology and Prevention of Vaccine-Preventable Diseases*. Public Health Foundation, 11th edition, 2011. ISBN 9780017066091.
- [3] B. Aulbach. *Gewöhnliche Differenzialgleichungen*. Spektrum Akademischer Verlag, 2004. ISBN 9783827414922.
- [4] E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280–7287, 2002. doi: 10.1073/pnas.082080899.
- [5] A. Briggs and M. Sculpher. An Introduction to Markov Modelling for Economic Evaluation. *PharmacoEconomics*, 13(4):397–409, 1998. doi: 10.2165/00019053-199813040-00003.
- [6] N. Britton. *Essential Mathematical Biology*. Springer Undergraduate Mathematics Series. Springer, 2003. ISBN 9781852335366.
- [7] J. Casti. *Would-be worlds: how simulation is changing the frontiers of science*. J. Wiley, 1997. ISBN 9780471123088.
- [8] Center for Infectious Disease Research & Policy. Avian Influenza (Bird Flu): Implications for Human Disease, 2011. URL [http://www.cidrap.umn.edu/cidrap/content/influenza/avianflu/biofacts/avflu\\_human.html](http://www.cidrap.umn.edu/cidrap/content/influenza/avianflu/biofacts/avflu_human.html).
- [9] C. Claes, R. R. Reinert, C. Vauth, and W. Greiner. Heptavalenter Pneumokokkenkonjugat-Impfstoff (PCV7), 2008. URL [http://gripsdb.dimdi.de/de/hta/hta\\_berichte/hta202\\_bericht\\_de.pdf](http://gripsdb.dimdi.de/de/hta/hta_berichte/hta202_bericht_de.pdf).
- [10] E. Codd. *Cellular automata*. Academic Press, 1968. ISBN 0121788504.
- [11] K. Dietz. Epidemiologic interference of virus populations. *Journal of Mathematical Biology*, 8(3):291–300, 1979. doi: 10.1007/BF00276314.

## BIBLIOGRAPHY

---

- [12] W. A. N. Dorland. *Dorland's Illustrated Medical Dictionary*. Saunders, 29th edition, 2000. ISBN 978-0721662541.
- [13] L. Esteva and C. Vargas. Coexistence of different serotypes of dengue virus. *Journal of Mathematical Biology*, 46(1):31–47, 2003. doi: 10.1007/s00285-002-0168-4.
- [14] P. E. M. Fine. Herd Immunity: History, Theory, Practice. *Epidemiologic Reviews*, 15(2):265–302, 1993.
- [15] M. A. Fuentes and M. N. Kuperman. Cellular automata and epidemiological models with spatial dependence. *Physica A: Statistical and Theoretical Physics*, 267(3):471 – 486, 1999.
- [16] H. Fuks and A. T. Lawniczak. Individual-based lattice model for spatial spread of epidemics. *Discrete Dynamics in Nature and Society*, 6(3):191–200, 2001. doi: 10.1155/S1026022601000206.
- [17] N. Ganguly, B. K. Sikdar, A. Deutsch, G. Canright, and P. P. Chaudhuri. A Survey on Cellular Automata. Technical report, Dresden University of Technology, 2003.
- [18] M. Gardner. The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223:120–123, 1970.
- [19] D. Greenberg. The Shifting Dynamics of Pneumococcal Invasive Disease after the Introduction of the Pneumococcal 7-Valent Conjugated Vaccine: Toward the New Pneumococcal Conjugated Vaccines. *Clinical Infectious Diseases*, 49(2):213–215, 2009. doi: 10.1086/599828.
- [20] S. Gupta, N. M. Ferguson, and R. M. Anderson. Vaccination and the population structure of antigenically diverse pathogens that exchange genetic material. *Proceedings of the Royal Society B: Biological Sciences*, 264(1387):1435–1443, 1997. doi: 10.1098/rspb.1997.0200.
- [21] M. Haber, A. Barskey, W. Baughman, L. Barker, C. G. Whitney, K. M. Shaw, W. Orenstein, and D. S. Stephens. Herd immunity and pneumococcal conjugate vaccine: A quantitative model. *Vaccine*, 25(29):5390 – 5398, 2007. doi: 10.1016/j.vaccine.2007.04.088.
- [22] K. K. Hsu, K. M. Shea, A. E. Stevenson, and S. I. Pelton. Changing Serotypes Causing Childhood Invasive Pneumococcal Disease: Massachusetts, 2001-2007. *The Pediatric Infectious Disease Journal*, 29(4):289–293, 2010. doi: 10.1097/INF.0b013e3181c15471.
- [23] D. J. Isaacman, D. R. Strutton, E. A. Kalpas, N. Horowicz-Mehler, L. S. Stern, R. Casciano, and V. Ciuryla. The impact of indirect (herd) protection on the cost-effectiveness of pneumococcal conjugate vaccine. *Clinical Therapeutics*, 30(2):341 – 357, 2008. doi: 10.1016/j.clinthera.2008.02.003.



- [24] N. R. Jennings and M. Wooldridge. On agent-based software engineering. *Artificial Intelligence*, 117(2):277–296, 2000. doi: 10.1016/S0004-3702(99)00107-1.
- [25] T. J. John and R. Samuel. Herd immunity and herd effect: new insights and definitions. *European Journal of Epidemiology*, 16(7):601–606, 2000. doi: 10.1023/A:1007626510002.
- [26] T. J. John, T. Cherian, M. C. Steinhoff, E. A. F. Simoes, and M. John. Etiology of Acute Respiratory Infections in Children in Tropical Southern India. *Reviews of Infectious Diseases*, 13(6):463–469, 1991. doi: 10.1093/clinids/13.Supplement\_6.S463.
- [27] G. Katriel and L. Stone. Pandemic Dynamics and the Breakdown of Herd Immunity. *PLoS ONE*, 5(3):e9565, 2010. doi: 10.1371/journal.pone.0009565.
- [28] W. O. Kermack and A. G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London. Series A*, 115(772):700–721, 1927. doi: 10.1098/rspa.1927.0118.
- [29] W. O. Kermack and A. G. McKendrick. Contributions to the Mathematical Theory of Epidemics. II. The Problem of Endemicity. *Proceedings of the Royal Society of London. Series A*, 138(834):55–83, 1932. doi: 10.1098/rspa.1932.0171.
- [30] W. O. Kermack and A. G. McKendrick. Contributions to the Mathematical Theory of Epidemics. III. Further Studies of the Problem of Endemicity. *Proceedings of the Royal Society of London. Series A*, 141(843):94–122, 1933. doi: 10.1098/rspa.1933.0106.
- [31] G. T. Keusch and K. J. Bart. Immunization Principles and Vaccine Use. In *Harrison’s Principles of Internal Medicine*, chapter 107. McGraw-Hill, 2008. ISBN 0071466339.
- [32] S. Ladhani, M. E. Ramsay, M. Chandra, and M. P. Slack. No evidence for *Haemophilus influenzae* serotype replacement in Europe after introduction of the Hib conjugate vaccine. *The Lancet Infectious Diseases*, 8(5):275 – 276, 2008. doi: 10.1016/S1473-3099(08)70078-1.
- [33] M. Lipsitch. Vaccination against colonizing bacteria with multiple serotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12):6571–6576, 1997.
- [34] M. Lipsitch. Bacterial Vaccines and Serotype Replacement: Lessons from *Haemophilus influenzae* and Prospects for *Streptococcus pneumoniae*. *Emerging Infectious Diseases*, 5(3):336–345, 1999. doi: 10.3201/eid0503.990304.
- [35] M. Lipsitch. Adaptive dynamics of infectious diseases: In pursuit of virulence management. chapter Vaccination and Serotype Replacement, pages 362–374. Cambridge University Press, 2002. ISBN 9780521781657.

## BIBLIOGRAPHY

---

- [36] E. S. Lysenko, R. S. Lijek, S. P. Brown, and J. N. Weiser. Within-Host Competition Drives Selection for the Capsule Virulence Determinant of *Streptococcus pneumoniae*. *Current biology*, 20(13):1222 – 1226, 2010. doi: 10.1016/j.cub.2010.05.051.
- [37] C. M. Macal and M. J. North. Tutorial on agent-based modeling and simulation. In *Proceedings of the 37th conference on Winter simulation*, WSC '05, pages 2–15. Winter Simulation Conference, 2005. ISBN 0-7803-9519-0.
- [38] C. M. Macal and M. J. North. Tutorial on agent-based modeling and simulation part 2: how to model with agents. In *Proceedings of the 38th conference on Winter simulation*, WSC '06, pages 73–83. Winter Simulation Conference, 2006. ISBN 1-4244-0501-7.
- [39] E. G. McIntosh, P. Conway, J. Willingham, and A. Lloyd. The cost-burden of paediatric pneumococcal disease in the UK and the potential cost-effectiveness of prevention using 7-valent pneumococcal conjugate vaccine. *Vaccine*, 21(19-20):2564 – 2572, 2003. doi: 10.1016/S0264-410X(03)00031-8.
- [40] M. North and C. Macal. *Managing business complexity: discovering strategic solutions with agent-based modeling and simulation*. Oxford University Press, 2007. ISBN 9780195172119.
- [41] S. Ohga, K. Okada, C. Miyazaki, K. Akazawa, and K. Ueda. The Measles Outbreak in Chikuhou District, Fukuoka, Japan, 1990: Correlation between Herd Immunity Level and Outbreak Size. *Pediatrics International*, 34(4):447–453, 1992. doi: 10.1111/j.1442-200X.1992.tb00985.x.
- [42] P. Ooi, K. G. S. Doraisingham, and A. Ling. Prevalence of varicella-zoster virus infection in Singapore. *Southeast Asian J Trop Med Public Health*, 23(1):22–25, 1992.
- [43] E. Parzen. *Stochastic Processes*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1987. ISBN 9780898714418.
- [44] A. Perisic and C. T. Bauch. Social Contact Networks and Disease Eradicability under Voluntary Vaccination. *PLoS Comput Biol*, 5(2):e1000280, 02 2009. doi: 10.1371/journal.pcbi.1000280.
- [45] L. R. Rabiner. Readings in speech recognition. chapter A tutorial on hidden Markov models and selected applications in speech recognition, pages 267–296. Morgan Kaufmann Publishers Inc., 1990. ISBN 1-55860-124-4.
- [46] P. Rooney and P. Coyle. The role of herd immunity in an epidemic cycle of hepatitis A. *Journal of Infection*, 24(3):327 – 331, 1992. doi: 10.1016/S0163-4453(05)80039-8.
- [47] S. Ross. *Stochastic processes*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley, 1983. ISBN 9780471099420.

- 
- [48] R. G. Sargent. Verification and validation of simulation models. In *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*, WSC '07, pages 124–137. IEEE Press, 2007. ISBN 1-4244-1306-0.
- [49] A. Scherer and A. McLean. Mathematical models of vaccination. *British Medical Bulletin*, 62(1):187–199, 2002. doi: 10.1093/bmb/62.1.187.
- [50] G. Schneckentreither, N. Popper, G. Zauner, and F. Breitenecker. Modelling SIR-type epidemics by ODEs, PDEs, difference equations and cellular automata - A comparative study. *Simulation Modelling Practice and Theory*, 16(8):1014 – 1023, 2008.
- [51] U. Siebert, O. Alagoz, A. M. Bayoumi, B. Jahn, D. K. Owens, D. J. Cohen, and K. M. Kuntz. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value in Health*, 15(6):812 – 820, 2012. doi: 10.1016/j.jval.2012.06.014.
- [52] H. Tijms. *A First Course in Stochastic Models*. John Wiley & Sons, 2003. ISBN 9780471498803.
- [53] W. W. C. Topley and G. S. Wilson. The Spread of Bacterial Infection. The Problem of Herd-Immunity. *Journal of Hygiene*, 21(3):243–249, 1923. doi: 10.1017/S0022172400031478.
- [54] C. Urach. *Modellierung und Simulation von Impfstrategien gegen Pneumokokken-erkrankungen: Markov- und Differentialgleichungsmodelle im Vergleich*. Master’s thesis, Vienna University of Technology, 2009.
- [55] S. Varadhan. *Stochastic Processes*. Courant Lecture Notes in Mathematics. Courant Institute of Mathematical Sciences, 2007. ISBN 9780821840856.
- [56] J. von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press Champaign, 1966.
- [57] D. A. Warrell, T. M. Cox, J. D. Firth, and E. J. J. M. Benz. *Oxford Textbook of Medicine*. Oxford Press, 4th edition, 2003. ISBN 0192629220.
- [58] J. P. Watt, L. J. Wolfson, K. L. O’Brien, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, O. S. Levine, R. Hajjeh, K. Mulholland, and T. Cherian. Burden of disease caused by Haemophilus influenzae type b in children younger than 5 years: global estimates. *The Lancet*, 374(9693):903 – 911, 2009. doi: 10.1016/S0140-6736(09)61203-4.
- [59] D. Wolf-Gladrow. *Lattice-gas cellular automata and lattice Boltzmann models*. Lecture notes in mathematics. Springer, 2000. ISBN 9783540669739.
- [60] S. Wolfram. *Theory and Applications of Cellular Automata*. World Scientific, 1986. ISBN 9971501236.

## BIBLIOGRAPHY

---

- [61] S. Wolfram. *A new kind of science*. Wolfram Media Inc., 2002. ISBN 1579550088.
- [62] M. Wooldridge. Agent-based software engineering. In *IEE Proceedings on Software Engineering*, pages 26–37, 1997.
- [63] S. Yakowitz, J. Gani, and R. Hayes. Cellular automaton modeling of epidemics. *Applied Mathematics and Computation*, 40(1):41 – 54, 1990. doi: 10.1016/0096-3003(90)90097-M.
- [64] Y. Zhang. *Mathematische Modelle zur Analyse der Koexistenz konkurrierender Serotypen*. PhD thesis, Eberhard Karls Universität Tübingen, 2006.
- [65] Štefan Emrich. *Comparison of mathematical models and development of a hybrid approach for the simulation and forecast of influenza epidemics within heterogeneous populations*. Master’s thesis, Vienna University of Technology, 2007.

# Lebenslauf

## ANGABEN ZUR PERSON

---

Name Florian Miksch  
Adresse Schröttergasse 3/14  
1100 Wien  
Telefon +43 650 5922269  
E-Mail fmiksch@osiris.tuwien.ac.at  
Staatsangehörigkeit Österreich  
Geburtsdatum 12. Juli 1984

## BERUFSERFAHRUNG

---

seit 06/2009 **dwh GmbH „dwh Simulation Services“**  
Tätigkeiten in den Bereichen Modellbildung & Simulation,  
Data-Mining, Projektmanagement, Lehrtätigkeit an der TU  
Wien, Betreuung von Bachelor- und Diplomarbeiten, Pub-  
likationen sowie Teilnahme an Konferenzen

06/2009 bis 12/2010 **Hauptverband der österreichischen Sozialver-  
sicherungsträger**  
Technische Unterstützung des KAL-Projekts („Katalog  
ambulanter Leistungen“) als externer Mitarbeiter

Sommermonate 2005,  
2006 und 2007 **Österreichisches Rotes Kreuz, Bezirksstelle Hart-  
berg**  
Arbeit als Rettungssanitäter und Einsatzfahrer

10/2002 bis 09/2003 **Zivildienst beim Österreichischen Roten Kreuz**  
Bezirksstelle Hartberg

## AUSBILDUNG

---

- seit 04/2009 **Doktoratsstudium der technischen Wissenschaften an der TU Wien**  
Dissertation über dynamische Modellbildung und Simulation von Epidemien
- 11/2007 bis 04/2008 **Auslandssemester an der University of the Philippines Diliman**  
Fakultät: College of Engineering  
Ort: Manila, Philippinen
- 10/2003 bis 03/2009 **Diplomstudium der technischen Mathematik an der TU Wien**  
Schwerpunkt: Mathematik in den Computerwissenschaften  
Abschluss mit ausgezeichnetem Erfolg
- 06/2002 **Matura am BG/BRG Hartberg**  
mit gutem Erfolg

## PERSÖNLICHE FÄHIGKEITEN UND KOMPETENZEN

---

- Muttersprache Deutsch
- Fremdsprache Englisch (verhandlungssicher)
- Führerschein A, B
- Programmierkenntnisse JAVA, C, MATLAB, objektorientierte Konzepte  
PHP, JavaScript  
SQL, Datenbankadministration