Contents

Chapter 1

Results

1.1 The Optimal Model

1.1.1 Times and Kernels

1.1.2 Model Size

Training data

The saved models are very large. A single Gaussian process model trained with 2000 training points is 31 MB large. Since one model is needed per process, this means that a relatively small model with a single expert with 2000 training points will require 36×31 MB ≈ 1.09 GB (endre dette tallet!). As seen from Table 1.2, the model size scales approximately as n^2, so the model size puts limits on the number of training points an optimal model should have.

Number of Features

There is little to no difference between models with 3 and 4 features.

Kernel

Using the Matérn kernel requires 29 bytes more than the RBF kernel, a number that does not change with the addition of training data.

How large is the optimal model?
Lagrede modeller:
dLdL, dLuL

$\tilde{q}_1\tilde{q}_2$	Time	C	$\ell_{\tilde{g}}$	$\ell_{\tilde{q}_1}$	$\ell_{\tilde{q}_2}$	$\ell_{\tilde{m}}$	α
$\tilde{u}_L\tilde{u}_L$	$\tilde{u}_L\tilde{d}_L$						

Table 1.1: Times and kernels for all processes.

Training data size	Model size [MB]
2000	31
3000	69
5000	191

Table 1.2: Size of saved GP models with 3 or 4 features and the Matérn kernel.

Figure 1.1: Mean relative deviance distributions for the processes $\tilde{u}_L\tilde{d}_L$, $\tilde{d}_L\tilde{d}_L$, $\tilde{s}_L\tilde{s}_L$ and $\tilde{c}_L\tilde{c}_L$ ($\tilde{u}_R\tilde{d}_R$, $\tilde{d}_R\tilde{d}_R$, $\tilde{s}_R\tilde{s}_R$ and $\tilde{c}_R\tilde{c}_R$).

1.2 Comparison with Prospino and NLL-fast

1.2.1 Mean Relative Deviance

Best Model for all Cross Sections

Compare Total with NLL-Fast

1.2.2 Cross Sections with Errors

- Plot masser mot hverandre, og plot sigma som funksjon av masser. Kanskje det holder med n masse? Evt med gjennomsnittsmassen?

PROBLEM: Det er aldri s stort sprik mellom dL og uL massen - kanskje funksjonen klikker pga det? mSUGRA kontra MSSM-24.

- DGP - Skalering av str med treningsdata (linear? n**2?) - Skalering av str med antall variable - Skalering av str med kernel

- Oevre grense fra hva man kan gi til andre (ikke stoerre enn 10GB) - Hvor stor er det den kan vre? - Nr er modellen saa stor at den er upraktisk?

Test: Kan man zippe en pickla GP? Hvor stor blir den da? sklearn?

- Mean relative deviance: Alle prosesser - Sammenlign med NLL-fast: Alle prosesser - Plot med NLL og Prospino

i tverrsnittdelen: dLdL og dRdR er like - Beregningene inneholder ingen EW-ledd, s den ser ikke forskjell paa R og L - Dette gjelder IKKE dersom EW NLO korreksjoner legges til - Ikke inkludert i prospino: qq -¿ X sq sq (ser forskjell paa R og L)

sjekk med dLuL og dRuR

— Hvor mange uavhengige prosesser finnes egentlig?

Aa gjoere:

Figure 1.2: Mean relative deviance distribution for the processes $\tilde{u}_L\tilde{d}_L$, $\tilde{u}_R\tilde{d}_R$, $\tilde{u}_L\tilde{d}_R$ and $\tilde{u}_R\tilde{d}_L$.

Figure 1.3: Mean relative deviance distribution for the processes $\tilde{u}_L\tilde{s}_L$, $\tilde{u}_R\tilde{s}_R$, $\tilde{u}_L\tilde{s}_R$ and $\tilde{u}_R\tilde{s}_L$.

Figure 1.4: Mean relative deviance distribution for the processes $\tilde{u}_L\tilde{c}_L$, $\tilde{u}_R\tilde{c}_R$, $\tilde{u}_L\tilde{c}_R$ and $\tilde{u}_R\tilde{c}_L$.

- Plotte sigmaer
- Hvor mange det reduseres til (tverrsnitt)
- Velg prosesser og sjekk skaleringer
- Lagre DGP-er
- Hvor stor er den ultimate modellen?
- Sjekk om man kan zippe en pickla GP, eller andre settinger
- Mean relative deviance for alle
- Sammenligne mrd med NLL-fast
- Plot mot NLL-fast og Prospino

1.2.3 Comparison with Prospino

A comparison with linspace-data from Prospino was attempted. Cross sections for the processes $\tilde{d}_L\tilde{d}_L$ and $\tilde{d}_L\tilde{u}_L$ were generated as a function of the mass $m_{\tilde{d}_L} = [200, 2400]$ GeV. All other masses, including $m_{\tilde{u}_L}$, were held at 1000 GeV. As can be seen in Fig. (1.8), the prediction for $\tilde{d}_L\tilde{d}_L$ is very good, and with good precision. The prediction for $\tilde{d}_L\tilde{u}_L$, however, is not good. The uncertainty is very big everywhere except at $m_{\tilde{d}_L} = 1000$ GeV. However, if the model trained on the process $\tilde{d}_R\tilde{u}_R$ is used, the prediction is very good and with little uncertainty, as seen in Fig. (1.9). This probably comes from the mass splittings of the MSSM. Left-handed squarks are part of the same $SU(2)_L$-doublet, and get their mass from the same parameter Q_1, with small mass splitting between same-generation squarks. The right-handed squarks, however, get their masses from different parameters $m^2_{\tilde{d}_1}$ and $m^2_{\tilde{u}_1}$, and so are not so bound. In the data set used here, the difference in $m_{\tilde{u}_L}$ and $m_{\tilde{d}_L}$ is very small, as seen in Fig. (1.10). For right-handed squarks, there is no apparent correlation, as they depend on different free parameters, as seen in Fig. (1.11).

This is why the $\tilde{d}_L\tilde{u}_L$-model predict so poorly for the points where $|m_{\tilde{d}_L} - m_{\tilde{u}_L}|$ is large, and the prediction is so good where $m_{\tilde{d}_L} = m_{\tilde{u}_L}$. As discussed in Sec. (SETT INN SEKSJON HVOR JEG DISKUTERER CROSS SECTIONS FOR R OG L), the form of the cross section is equal for equal-chirality final-squarks,

Figure 1.5: Mean relative deviance distribution for the processes $\tilde{d}_L\tilde{s}_L$, $\tilde{d}_R\tilde{s}_R$, $\tilde{d}_L\tilde{s}_R$ and $\tilde{d}_R\tilde{s}_L$.

Figure 1.6: Mean relative deviance distribution for the processes $d_L c_L$, $d_R c_R$, $d_L c_R$ and $d_R c_L$.

Figure 1.7: Mean relative deviance distribution for the processes $s_L c_L$, $s_R c_R$, $s_L c_R$ and $s_R c_L$.

so the model trained on $\tilde{d}_R \tilde{u}_R$ is able to do a good prediction for $\tilde{d}_L \tilde{u}_L$, and since the right-handed masses depend on different parameters it allows for large mass differences. The same applies for the second generation squarks.

For the comparison with `Prospino 2.1` 10 experts with 5000 training points were used, and $m_{\tilde{d}_L} = m_{\tilde{u}_L}$. This is an approximation, since, as previously mentioned, the mass splitting is in reality given by

$$m^2_{\tilde{d}_L} - m^2_{\tilde{u}_L} = -\cos(2\beta) m^2_W. \tag{1.1}$$

As $\cos 2\beta$ is relatively small in the given data set, $\tan\beta \approx 11.697$, the masses are set to be equal.

Figure 1.8: Cross sections for $\tilde{d}_L \tilde{d}_L$ and $\tilde{u}_L \tilde{d}_L$, using $m_{\tilde{d}_L} = [200, 2400]$,GeV and all other masses set to $m_i = 1000$ GeV generated by prospino (crosses) and predicted by the GP (lines with errors). The GP models used are for $\tilde{d}_L \tilde{d}_L$ and $\tilde{d}_L \tilde{u}_L$.

Figure 1.9: Cross sections for $\tilde{d}_L\tilde{d}_L$ and $\tilde{u}_L\tilde{d}_L$, using $m_{\tilde{d}_L} = [200, 2400]$,GeV and all other masses set to $m_i = 1000$ GeV generated by prospino (crosses) and predicted by the GP (lines with errors). The GP models used are for $\tilde{d}_L\tilde{d}_L$ and $\tilde{d}_R\tilde{u}_R$.

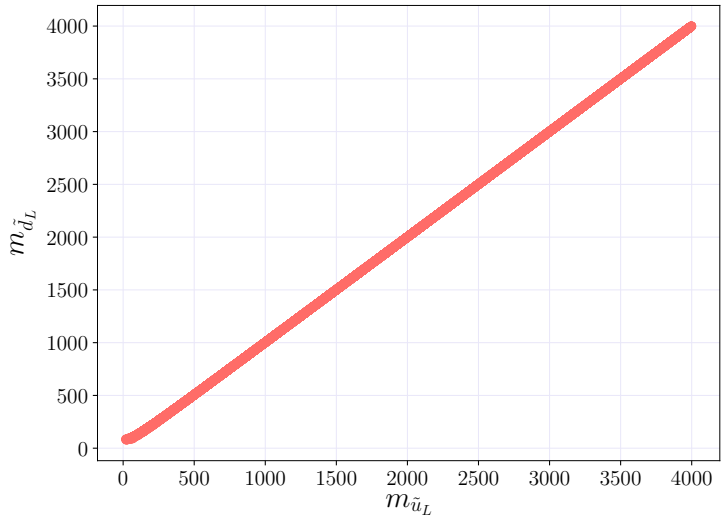

Figure 1.10: Scatter plot of the left-handed squark masses $m_{\tilde{u}_L}$ and $m_{\tilde{d}_L}$. They are very strongly correlated.

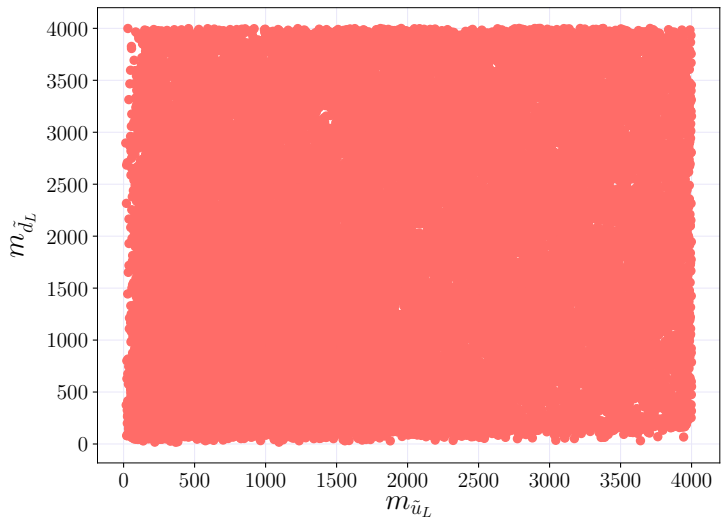

Figure 1.11: Scatter plot of the left-handed squark masses $m_{\tilde{u}_R}$ and $m_{\tilde{d}_R}$. They are not correlated.