Prospects for the measurement of $\sigma_H \times BR(H \rightarrow \mu\mu)$ at 3 TeV

Alessandro Montella, Massimo Casarsa, Vieri Candelise
University of Trieste and INFN-Trieste

June 22, 2021
Motivations

➢ The measurement of the Higgs boson branching ratios provides a strong test of the Standard Model and possible new physics. As a consequence, the exploration of the Higgs sector is going to be an essential part of the precision physics programme at the future colliders.

➢ Only the couplings of the Higgs boson to the vector bosons $W$ and $Z$ and the third-generation charged fermions of the Standard Model have been measured to date, while the coupling to the first and second-generation fermions are yet to be confirmed.

➢ The $H \rightarrow \mu \mu$ decay provides a sensitive probe for a precise measurement of the coupling of the Higgs boson to the second generation fermions at the muon collider.
Analysis strategy

This is the first approach to the analysis of the $H \rightarrow \mu\mu$ decay at a muon collider and the strategy adopted was inspired by the CLIC's strategy

- A loose preselection is applied to the events with a pair of opposite-charge reconstructed muons with an invariant mass in the range of $\pm 20$ GeV around the nominal mass of the Higgs boson.

- A multivariate event classifier based on a boosted decision tree (BDT) is trained to discriminate between signal and background preselected events.

- The event selection is performed with a cut on the BDT distributions.

- The signal is extracted with a fit to the di-muon system invariant mass distribution of the selected events.
Monte Carlo samples

- Events are generated using **MADGRAPH5_aMC@NLO v3.1.0** (hard events) + **Pythia8** (parton showering, hadronization and decays) and fully simulated and reconstructed with the package **MuonColliderSoft**.

<table>
<thead>
<tr>
<th>Process</th>
<th>Kinematic requirements</th>
<th>$\sigma$ (fb)</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^+\mu^- \rightarrow H\nu_\mu \bar{\nu}_\mu, \ H \rightarrow \mu^+\mu^-$</td>
<td>$m_{\mu\mu} &gt; 100$ GeV</td>
<td>0.10940 ± 0.00027</td>
<td>20000</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow H\mu^+\mu^-, \ H \rightarrow \mu^+\mu^-$</td>
<td>$m_{\mu\mu} &gt; 100$ GeV</td>
<td>0.009811 ± 0.000029</td>
<td>20000</td>
</tr>
</tbody>
</table>
| $\mu^+\mu^- \rightarrow \mu^+\mu^-\nu\bar{\nu}_\mu$                     | $100 < m_{\mu\mu} < 150$ GeV,  
$8^\circ < \theta_\mu < 172^\circ$ GeV | 11.0920 ± 0.0078         | 255000 |
| $\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$                             | $m_{\mu+\mu^-} > 100$ GeV                       | 1530.50 ± 0.56          | 2000000|
| $\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$                             | $100 < m_{\mu^+\mu^-} < 150$ GeV,  
$8^\circ < \theta_\mu < 172^\circ$ GeV | 297.400 ± 0.074          | 1785000|
| $\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$                             | $100 < m_{\mu^+\mu^-} < 150$ GeV,  
$8^\circ < \theta_\mu < 172^\circ$ GeV | $75.580 \pm 0.023$       | 128000 |
| $\mu^+\mu^- \rightarrow tt \rightarrow W^+W^-bb, W^\pm \rightarrow \mu^\pm \nu_\mu (\bar{\nu}_\mu)$ | $m_{\mu^+\mu^-} > 50$ GeV                        | 0.31950 ± 0.00092       | 10000  |

$\sqrt{s} = 3$ TeV, $m_H = 125$ GeV, $\Gamma_H = 4.07 \times 10^{-3}$ GeV
A preselection is applied to the samples before the event classification and the final selection.

---

**Preselection requirements**

- Two opposite-charge muons
- $10^\circ < \theta_{\mu} < 170^\circ$
- $105 < m_{\mu^+\mu^-} < 145$ GeV
- $p_T(\mu^\pm) > 5$ GeV
- $p_T(\mu^+\mu^-) > 30$ GeV
- $p_T(\mu^+) + p_T(\mu^-) > 50$ GeV

---

The requirements on $p_T(\mu^+\mu^-)$ and $p_T(\mu^+) + p_T(\mu^-)$ are aimed to reduce the background contribution in the regions where (almost) no signal is present.
**Expected number of events**

- All the samples are normalized assuming an integrated luminosity of $L_{\text{int}} = 1\text{ ab}^{-1}$.

<table>
<thead>
<tr>
<th>Process</th>
<th>Expected events with $105 &lt; m_{\mu\mu} &lt; 145\text{ GeV}$ (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^+\mu^- \rightarrow H\nu_\mu\bar{\nu}_\mu$, $H \rightarrow \mu^+\mu^-$</td>
<td>76.5</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow H\mu^+\mu^-$, $H \rightarrow \mu^+\mu^-$</td>
<td>6.4</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow \mu^+\mu^-\nu\bar{\nu}_\mu$</td>
<td>5027.9</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$</td>
<td>34729.7</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow t\bar{t} \rightarrow W^+W^-b\bar{b}$, $W^\pm \rightarrow \mu^\pm\nu_\mu(\bar{\nu}_\mu)$</td>
<td>10.8</td>
</tr>
</tbody>
</table>

(*) After preselection.

- After the preselection, about 0.2% of the events in the chosen invariant mass range are expected to be signal events.
Event classification

- The event classification is performed using a multivariate classifier based on a boosted decision tree (BDT), implemented in the TMVA toolkit for multivariate analysis with ROOT.

- Two BDT-based classifiers are trained separately to discriminate between the signal and two main background contributions:
  - **BDT 1**: SIGNAL vs $\mu^+\mu^- \rightarrow \mu^+\mu^-\nu\bar{\nu}$
  - **BDT 2**: SIGNAL vs $\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$

- Half of the events of the signal and the background samples are reserved for the training and testing of the BDT-based classifiers and are not used for the subsequent event selection and analysis.

**OBSERVABLES USED FOR THE CLASSIFICATION**

- Transverse momentum of the di-muon system
- Modulus of the 3-momentum of the di-muon system
- Polar angle of the di-muon system
- Scalar sum of the transverse momentum of the two selected muons
- Total visible energy of the event
- Visible energy of the event excluding the two selected muons
- Recoil mass
- Missing transverse energy
- Boost of the di-muon system
- Cosine of the helicity angle (with respect to the negative-charge muon)
- Cosine of the angle between the two selected muons
- Angular separation of the two selected muons
Distribution of $p_T$, $|\vec{p}|$ and $\theta$

- The two signal contributions present a very similar shape of the distributions with respect to the background contributions.
Distribution of other observables

Event total visible energy

Event visible energy excluding the di-muon system

Recoil mass, defined as

\[ M_{\text{recoil}} = \sqrt{(p_\mu^+ + p_\mu^- - p_{\mu\mu})^2} \]

Scalar sum of the two selected muons \( p_T \)

Missing transverse energy

\[ E_T^{\text{miss}} = -\sum_i p_T(i) \]
Distribution of other observables

Boost of the di-muon system $\beta_{\mu \mu} = \frac{p_{\mu \mu}}{E_{\mu \mu}}$

Cosine of the helicity angle $\cos \theta^* = \frac{(\vec{p}_{\mu^{-}})(\vec{p}_{\mu \mu})}{(p_{\mu^{-}})(p_{\mu \mu})}$

Cosine of the angle between the 3-vectors of the two selected muons

Angular separation of the two selected muons
BDT training

**SIGNAL vs**

\( \mu^+ \mu^- \rightarrow \mu^+ \mu^- \nu \bar{\nu} \)

(BDT1)

**SIGNAL vs**

\( \mu^+ \mu^- \rightarrow \mu^+ \mu^- \mu^+ \mu^- \)

(BDT2)
Event selection

➢ The BDT scores are evaluated for each sample reserved for the analysis using the two trained classifiers separately.

➢ The event selection is performed with the application of a cut on the values of both the BDT classifiers, such that an event is classified as a "signal event" if it is successfully discriminated against at least one of the two backgrounds considered for the training.

Stacks of the BDT value distributions

➢ The classifiers cut values are optimized simultaneously to maximize the significance, defined as $S/\sqrt{S + B}$, where $S$ and $B$ are the number of the selected signal and background events respectively.

Best cut $= -0.34$  
Best cut $= -0.22$

Significance at best cuts $S/\sqrt{S + B} = 0.76$
Event selection

Expected events after the selection

<table>
<thead>
<tr>
<th>Process</th>
<th>Expected events with $105 &lt; m_{\mu\mu} &lt; 145$ GeV (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^+\mu^- \rightarrow H\nu_\mu\bar{\nu}_\mu, \nu\rightarrow \mu^+\mu^-$</td>
<td>24.2</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow H\mu^+\mu^-, \nu\rightarrow \mu^+\mu^-$</td>
<td>1.6</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow \mu^+\mu^-\nu\bar{\nu}_\mu$</td>
<td>636.5</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$</td>
<td>476.4</td>
</tr>
<tr>
<td>$\mu^+\mu^- \rightarrow t\bar{t} \rightarrow W^+W^-b\bar{b}, W^\pm \rightarrow \mu^\pm\nu_\mu(\bar{\nu}_\mu)$</td>
<td>1.1</td>
</tr>
</tbody>
</table>

(*) After final selection.
Signal extraction with a fit to the $m_{\mu\mu}$ distribution

- The quantity $\sigma_H \times BR(H \to \mu\mu)$ is determined from:

\[
\sigma_H \times BR(H \to \mu\mu) = \frac{N_S}{L_{int} \epsilon_s}
\]

  - $N_S$ = number of signal events
  - $L_{int}$ = integrated luminosity of the experiment
  - $\epsilon_s$ = total signal counting efficiency

The statistical uncertainty is dominated by the statistical uncertainty $\sigma_{N_S}$ on the number of signal events $N_S$.

- In a real experiment, the number of signal events would be estimated with a fit to the invariant mass distribution of the di-muon system.

- In our case, the fit to the invariant mass distribution is used to estimate the sensitivity on $N_S$ with a toy-MC study, given the event yields expected in 1 $ab^{-1}$. 
Di-muon invariant mass fit

- Using the RooFit toolkit for data modeling, an unbinned extended maximum likelihood (EML) fit is performed to the di-muon system invariant mass distribution to estimate the uncertainty on $N_S$.

- The unbinned EML fit is performed with a function $f(m_{\mu\mu}) = N_S \cdot f_S(m_{\mu\mu}) + N_B \cdot f_B(m_{\mu\mu})$, where $f_S$ and $f_B$ are the probability density functions that describe the distribution of the invariant mass of the signal and the background events, respectively, and $N_S$ and $N_B$ are the free parameters estimated with the EML fit.

- The probability density functions $f_S$ and $f_B$ are extracted with an unbinned maximum likelihood fit to the di-muon system invariant mass of the selected signal and background events.
The toy-MC study is performed by generating 10000 pseudo-experiments.

The pseudo-data for each one of the 10000 pseudo-experiments are obtained with the random generation of $N_S'$ signal and $N_B'$ background $m_{\mu\mu}$ values from their previously extracted probability density functions $f_S$ and $f_B$.

The number of signal and background events generated for each pseudo-experiment is obtained from a Poisson distribution with a mean value of $E[N_S'] = 25.8$ for the signal and with a mean value of $E[N_B'] = 1114$ for the background.

For each pseudo-experiment, an unbinned EML fit is performed to the invariant mass distribution with the function:

$$ f(m_{\mu\mu}) = N_S' \cdot f_S(m_{\mu\mu}) + N_B' \cdot f_B(m_{\mu\mu}) $$
Sensitivity of $N_S$

➢ The RMS of the signal count distribution of the 10000 pseudo-experiments is taken as the estimate of $\sigma_{N_S}$

\[
\sigma_{N_S} = 25.8 \pm 9.9
\]

\[
\frac{\sigma_{N_S}}{\langle N_S \rangle} = 38\%
\]

$L_{\text{int}} = 1 \text{ ab}^{-1}$
Conclusions

➢ For a 3 TeV muon collider, the number of $H \rightarrow \mu\mu$ events in the range $105 < m_{\mu\mu} < 145$ GeV is expected to be $\langle N_S \rangle = 25.8 \pm 9.9$ for a dataset corresponding to an integrated luminosity of 1 ab$^{-1}$.

➢ The relative uncertainty on $\langle N_S \rangle$, which corresponds to the statistical uncertainty of the measurement, is estimated to be 38%.

➢ By comparison, for a 3 TeV CLIC collider with a dataset corresponding to an integrated luminosity of 2 ab$^{-1}$, the number of signal events in the same $m_{\mu\mu}$ range is expected to be about 53 with a relative uncertainty of 23% without the $\gamma\gamma \rightarrow$ hadrons background (with the inclusion of this background contribution, the relative uncertainty is esitmated to be 26.3%). (*)

With a rough scaling of our result to a 2 ab$^{-1}$ statistics, we obtain a sensitivity of about 27%, which shows that the performances for this measurement at the muon collider and at the CLIC collider would be similar, which is a good result considering that in the CLIC case there is no equivalent of the $\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$ background.

Future improvements to the analysis

➢ The analysis has been performed without taking into account the effects of the beam-induced background, which cannot be neglected in the context of a muon collider.

➢ A better strategy for the suppression of the $\mu^+\mu^- \rightarrow \mu^+\mu^-\mu^+\mu^-$ background would improve the performances with respect to CLIC.

➢ This study has been performed considering a 3 TeV muon collider but we expect increased signal yields and an improved sensitivity at higher energies.
THANK YOU!
BACKUP
Estimate of $\varepsilon_S$

- Preselection efficiency = 95%
- MVA selection efficiency = 31%

$\varepsilon_S = 30\%$
Distribution of $N_B$

Distribution of $N_B$

Distribution of $N_B$ pulls

RMS = 34.38 ± 0.24
Mean = 1114.03 ± 0.34
Entries = 10000

pullMean = -0.01742 ± 0.0100
pullSigma = 0.9986 ± 0.0071
Tracking efficiency as a function of $\theta$ and $p_T$

➢ Single muons without BIB

![Graph showing tracking efficiency as a function of $\theta$ and $p_T$.]
Resolution in $p_T$ as a function of $\theta$ and $p_T$

- Single muons without BIB

![Graph showing resolution in $p_T$ as a function of $\theta$ and $p_T$.](image-url)
Linear correlation coefficients (BDT training 1)
Linear correlation coefficients (BDT training 2)

**Correlation Matrix (signal)**

- **Linear correlation coefficients in %**

<table>
<thead>
<tr>
<th></th>
<th>[p] of the di-muon system</th>
<th>Event total visible energy</th>
<th>Missing Transverse Energy</th>
<th>Other visible energy</th>
<th>Cosine between the two muons</th>
<th>DeltaR between the two muons</th>
<th>Recoil mass</th>
<th>Cosine helicity angle</th>
<th>Di-muon system polar angle</th>
<th>Di-muon system boost</th>
<th>Single-muon pT scalar sum</th>
<th>Di-muon system pT</th>
</tr>
</thead>
<tbody>
<tr>
<td>[p] of the di-muon system</td>
<td>36 37 46 -9 -61 -64 -15 46 11 37 64 100</td>
<td>56 57 64 1 -12 23 69 26 58 100 64</td>
<td>-97 -99 -44 1 -3 60 36 -46 -10 100 -58 -37</td>
<td>3 3 10 -1 -12 -2 9 100 10 26 11</td>
<td>45 46 83 1 -65 -33 100 9 46 69 45</td>
<td>40 43 -26 2 -1 22 100 33 -2 36 -23 15</td>
<td>-58 -60 -64 -1 100 22 -65 -12 60 -97 -64</td>
<td>2 3 1 100 1 1 -3 1 1 1</td>
<td>-1 -1 1 100 1 2 1 1 -3 1 0</td>
<td>-47 -44 100 1 -64 -26 83 10 44 69 45</td>
<td>97 100 44 -1 3 -60 -36 46 3 99 57 37</td>
<td>100 97 47 -1 2 -58 -40 45 3 97 56 36</td>
</tr>
</tbody>
</table>

**Correlation Matrix (background)**

- **Linear correlation coefficients in %**

<table>
<thead>
<tr>
<th></th>
<th>[p] of the di-muon system</th>
<th>Event total visible energy</th>
<th>Missing Transverse Energy</th>
<th>Other visible energy</th>
<th>Cosine between the two muons</th>
<th>DeltaR between the two muons</th>
<th>Recoil mass</th>
<th>Cosine helicity angle</th>
<th>Di-muon system polar angle</th>
<th>Di-muon system boost</th>
<th>Single-muon pT scalar sum</th>
<th>Di-muon system pT</th>
</tr>
</thead>
<tbody>
<tr>
<td>[p] of the di-muon system</td>
<td>20 23 42 16 -67 -58 -6 38 9 23 56 100</td>
<td>25 41 71 0 96 -11 63 38 46 100 98</td>
<td>-60 -97 35 1 41 20 -50 -30 100 -46 -23</td>
<td>10 10 9 -16 -2 9 100 30 38 9</td>
<td>26 51 68 1 -64 -25 100 9 -50 63 38</td>
<td>28 20 -15 1 110 100 -25 -2 20 -11 -6</td>
<td>-35 -41 -71 1 100 11 -64 -15 11 -96 -58</td>
<td>-18 100 0 0 0 0 0 0 0 0 0</td>
<td>-180 -180 0 0 0 0 0 0 0 0 0</td>
<td>-38 -35 100 1 -71 -15 68 9 38 71 44</td>
<td>38 50 35 0 -41 -20 51 10 -87 -41 23</td>
<td>72 100 35 1 -41 -20 51 10 -87 -41 23</td>
</tr>
</tbody>
</table>