Common

Data

```
data = {{0.34601851851851845`, 9.29045864275833`}, {0.3834259259259259`, 7.777297227353334`},
   {0.42083333333333334`, 6.475978410105039`}, {0.626574074074074`, 4.902290538083845`},
   {0.47694444444444445`, 3.5404452642193505`}, {0.6452777777777776`, 1.6943883374252575`},
   {2.010648148148148`, 1.1799134561875597`}, {3.020648148148148`, 1.210176684495659`},
   {4.011944444444444`, 1.02859731464706`}, {5.9571296296296286`, 1.0891237712632593`},
   {7.453425925925925`, 1.1799134561875597`}, {8.762685185185184`, 1.210176684495659`},
   {9.828796296296296`, 1.7246515657333572`}, {9.941018518518518`, 3.7522878623760496`},
   {9.772685185185184`, 5.840450615634942`}, {9.8475`, 7.807560455661434`},
   {9.75398148148148`, 9.865459980612226`}, {8.18287037037037`, 9.502301240915028`},
   {6.144166666666665`, 9.593090925839327`}, {4.7413888888888875`, 9.623354154147428`},
   {3.1515740740740736`, 9.804933523996027`}, {1.580462962962963`, 9.472038012606928`},
   {3.2450925925925924`, 8.079929510434333`}, {3.319907407407407`, 7.383875259348036`},
   {3.3573148148148144`, 6.839137149802238`}, {3.3012037037037034`, 6.475978410105039`},
   {3.263796296296296`, 5.931240300559241`}, {3.263796296296296`, 5.386502191013443`},
   {3.2263888888888888`, 4.962816994700045`}, {3.2263888888888888`, 4.327289200229948`},
   {3.2263888888888888`, 3.7522878623760496`}, {3.1702777777777773`, 3.086496839597852`},
   {3.13287037037037`, 2.4509690451277546`}, {0.626574074074074`, 7.958876597201934`},
   {0.6639814814814815`, 7.020716519650837`}, {0.6452777777777776`, 6.173346127024041`},
   {0.6452777777777776`, 5.295712506089144`}, {0.626574074074074`, 4.599658255002846`},
   {0.5891666666666666`, 3.5404452642193505`}, {0.5891666666666666`, 2.7838645565168534`},
   {0.5891666666666666`, 2.057547077122456`}, {0.626574074074074`, 8.352298565207231`},
   {0.47694444444444445`, 7.353612031039936`}, {0.3273148148148148`, 3.4799188076031506`},
   {0.34601851851851845`, 5.325975734397244`}, {0.439537037037037`, 4.448342113462347`},
   {0.3647222222222222`, 3.358865894370751`}, {0.42083333333333334`, 2.1785999903548556`},
   {1.4121296296296295`, 1.3009663694199594`}, {0.7949074074074074`, 9.562827697531228`},
   {2.440833333333333`, 9.744407067379827`}, {2.833611111111111`, 9.13914250121783`},
   {2.4221296296296293`, 9.472038012606928`}, {1.4121296296296295`, 9.683880610763627`},
   {1.991944444444444`, 9.804933523996027`}, {0.9632407407407406`, 0.9983340863389598`},
   {1.6178703703703703`, 0.8167547164903608`}, {2.515648148148148`, 0.8167547164903608`},
   {2.7400925925925925`, 1.2707031411118592`}, {1.7862037037037037`, 7.535191400888536`},
   {1.7862037037037037`, 5.265449277781044`}, {1.6739814814814813`, 3.5404452642193505`}};
```

```
σ = 1;
```

$$K[x\_] := \frac{1}{2 * \pi * \sigma^2} e^{-\frac{x[[1]]^2 + x[[2]]^2}{2 * \sigma^2}}$$

```
h = 0.8;
```

$$error[w1\_, w2\_] := \frac{1}{h^2 * Length[data]} \sum_{i=1}^{Length[data]} K\left[\frac{data[[i]] - \{w1, w2\}}{h}\right]$$

```
grad1 = D[error[w1, w2], {w1}];
grad2 = D[error[w1, w2], {w2}];

(*errorData=<||>;
errorData["w1"]=Table[w1,{w1,0,9,0.1}];
errorData["w2"]=Table[w2,{w2,1,8,0.1}];
errorData["error"]=Table[error[w1,w2],{w2,1,8,0.1},{w1,0,9,0.1}];
Export[FileNameJoin[{NotebookDirectory[],"AdamOptimizer_FirstOrderMomentsFunctionData.json"}],
 errorData,"Compact"→True];*)

w1Start = 0.8;
w2Start = 3.6;
```

```
clear[] := Block[{},
  (* This removes all subscript values
   (https://mathematica.stackexchange.com/a/38650/41589) *)
  Clear["Subscript"];

  grad1[w1_, w2_] = grad1;
  grad2[w1_, w2_] = grad2;
 ]
```

$$w_{i\_}[t\_] := w_i[t] = w_i[t-1] - \text{eta} * \frac{\frac{m_i[t]}{1-\text{beta1}^t}}{\sqrt{\frac{s_i[t]}{1-\text{beta2}^t}} + 1};$$

```
adam[beta1_, beta2_, eta_, iterations_: 1000] := Block[{},
  clear[];

  s_i_[0] := 0;
  m_i_[0] := 0;
  w1[0] := w1Start;
  w2[0] := w2Start;

  m_i_[t_] := m_i[t] = beta1 * m_i[t - 1] + (1 - beta1) * grad_i[w1[t - 1], w2[t - 1]];
  s_i_[t_] := s_i[t] = beta2 * s_i[t - 1] + (1 - beta2) * grad_i[w1[t - 1], w2[t - 1]]^2;
  w_i_[t_] := w_i[t] = w_i[t - 1] - eta * m_i[t] / (1 - beta1^t);

  Table[{w1[t], w2[t], error[w1[t], w2[t]]}, {t, 0, 1000}]
 ]
```

Plots

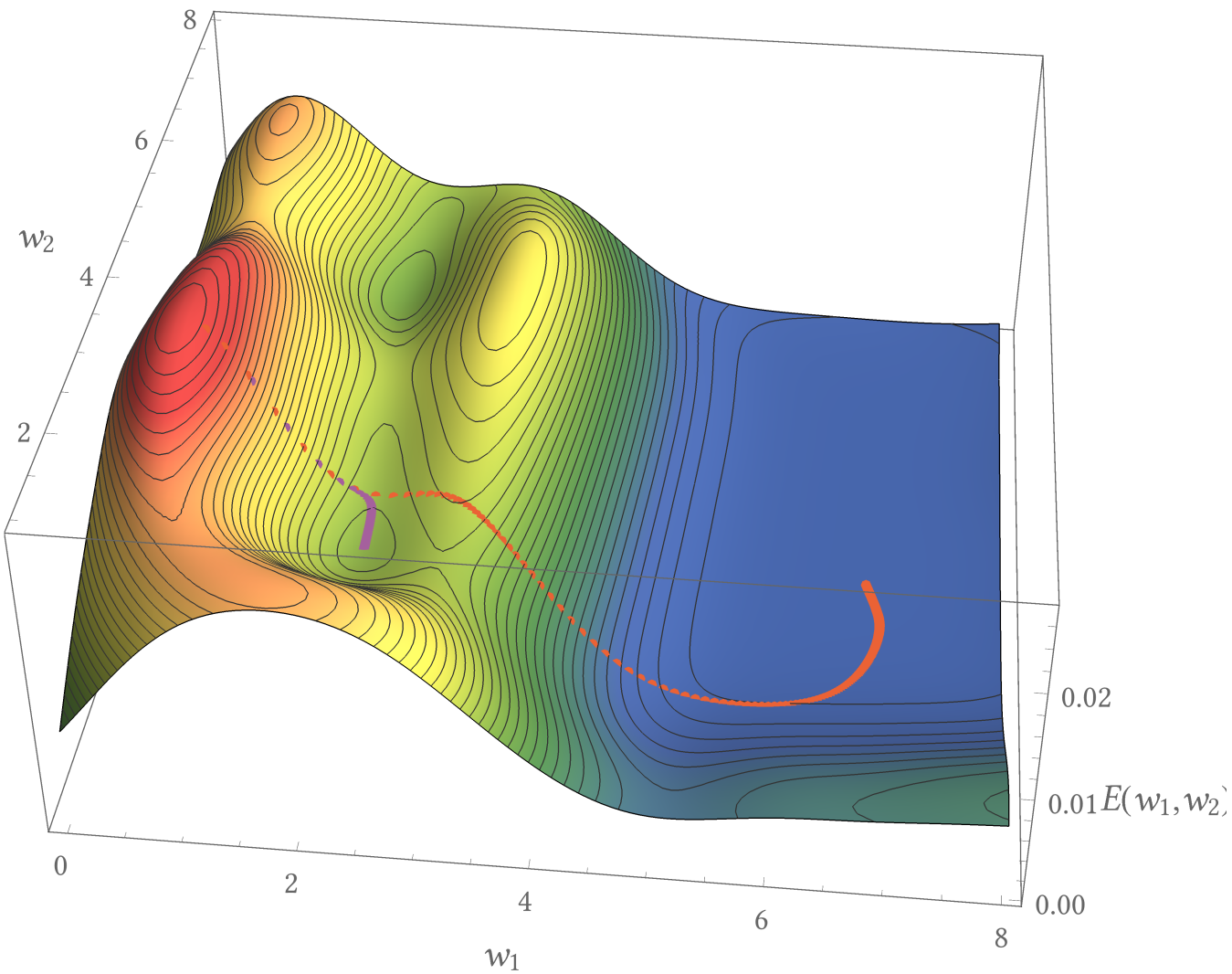
Trajectory

```
color1 = ■;
color2 = ■;

vp = Options[Graphics3D, ViewPoint][[1, 2]];

vp
```

{1.3, -2.4, 2.}

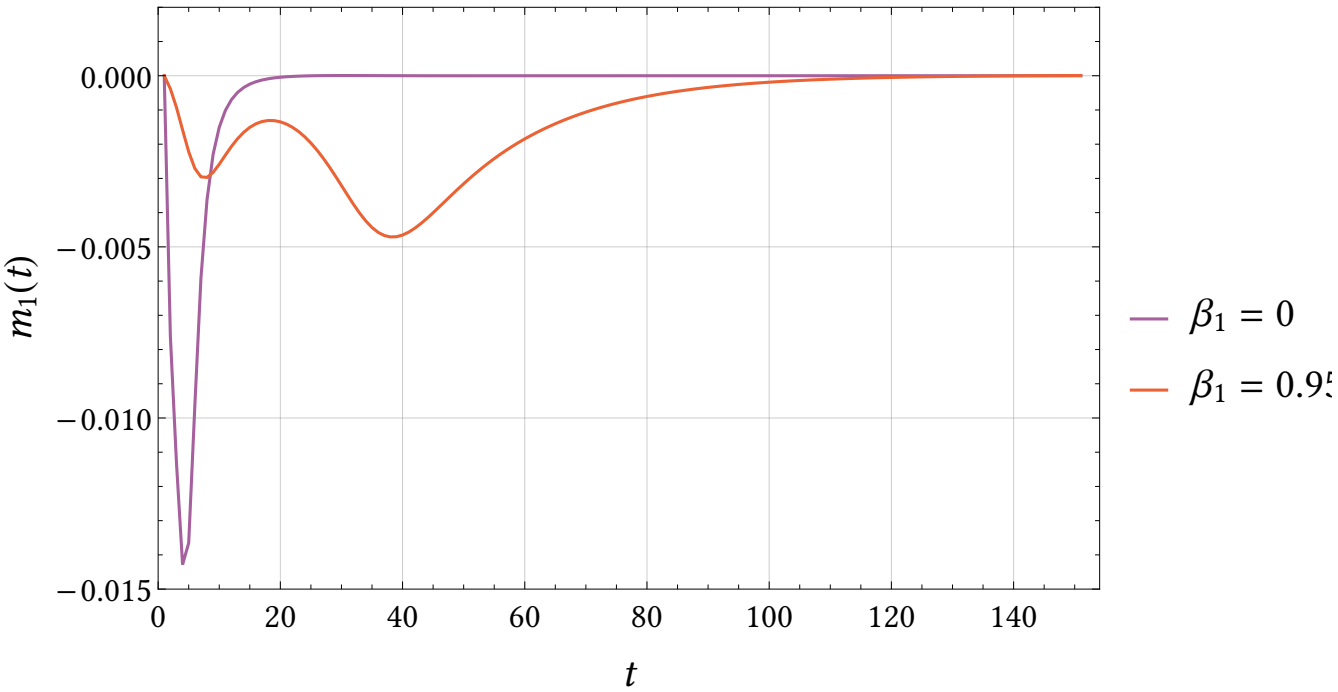
```
trajectoryClassic = adam[0, 1, 20];
m1Classic = Table[m1[t], {t, 0, 150}];
trajectoryAdam = adam[0.95, 1, 20];
m1Adam = Table[m1[t], {t, 0, 150}];
```

```
plotSurface = Show[
  Plot3D[error[w1, w2], {w1, 0, 8}, {w2, 1, 8},
   PlotTheme → "myTheme",
   AxesLabel → {"w1", "w2", Row[{it["E"], "(w1,w2)"}]},
   ImageSize → 700,
   PlotStyle → Opacity[0.7],
   ColorFunction → (ColorData["DarkRainbow"][#3] &),
   PlotLegends → Automatic,
   MeshFunctions -> {#3 &},
   Mesh → 40,
   PlotPoints → 100,
   MaxRecursion → 10,
   ViewPoint → {0.2828737948779217`, -2.6762223452154776`, 2.051296267032315`},
   PerformanceGoal → "Quality"
  ],
  ListPointPlot3D[{trajectoryClassic, trajectoryAdam},
   PlotTheme → "myTheme",
   PlotStyle → {Directive[color1, PointSize[0.01]], Directive[PointSize[0.01], color2]},
   PlotLegends → PointLegend[{color1, color2}, {"β1 = 0", "β1 = 0.9"}, LegendMarkerSize → 20]
  ]
 ]
```



```
(*β1Min=0;
β1Max=0.99;
β1Step=0.01;
Do[
 Export[
   FileNameJoin[{
     NotebookDirectory[],
     "beta1="<>IntegerString[Round[(β1-β1Min)/β1Step,1],10,2]<>".json"
    }],
   adam[β1,1,20,500]ᵀ,
   "Compact"→True
  ];
 ,{β1,β1Min,β1Max,β1Step}]*)
```

```
plotUpdate = ListLinePlot[{m1Classic, m1Adam},
  PlotTheme → "myTheme",
  PlotRange → {{0, Automatic}, {-0.015, 0.002}},
  PlotStyle → {color1, color2},
  FrameLabel → {it["t"], Row[{Subscript[it["m"], "1"], "(", it["t"], ")"}]},
  PlotLegends → {"β₁ = 0", "β₁ = 0.95"},
  GridLines → Automatic,
  Frame → True,
  Axes → False
 ]
```

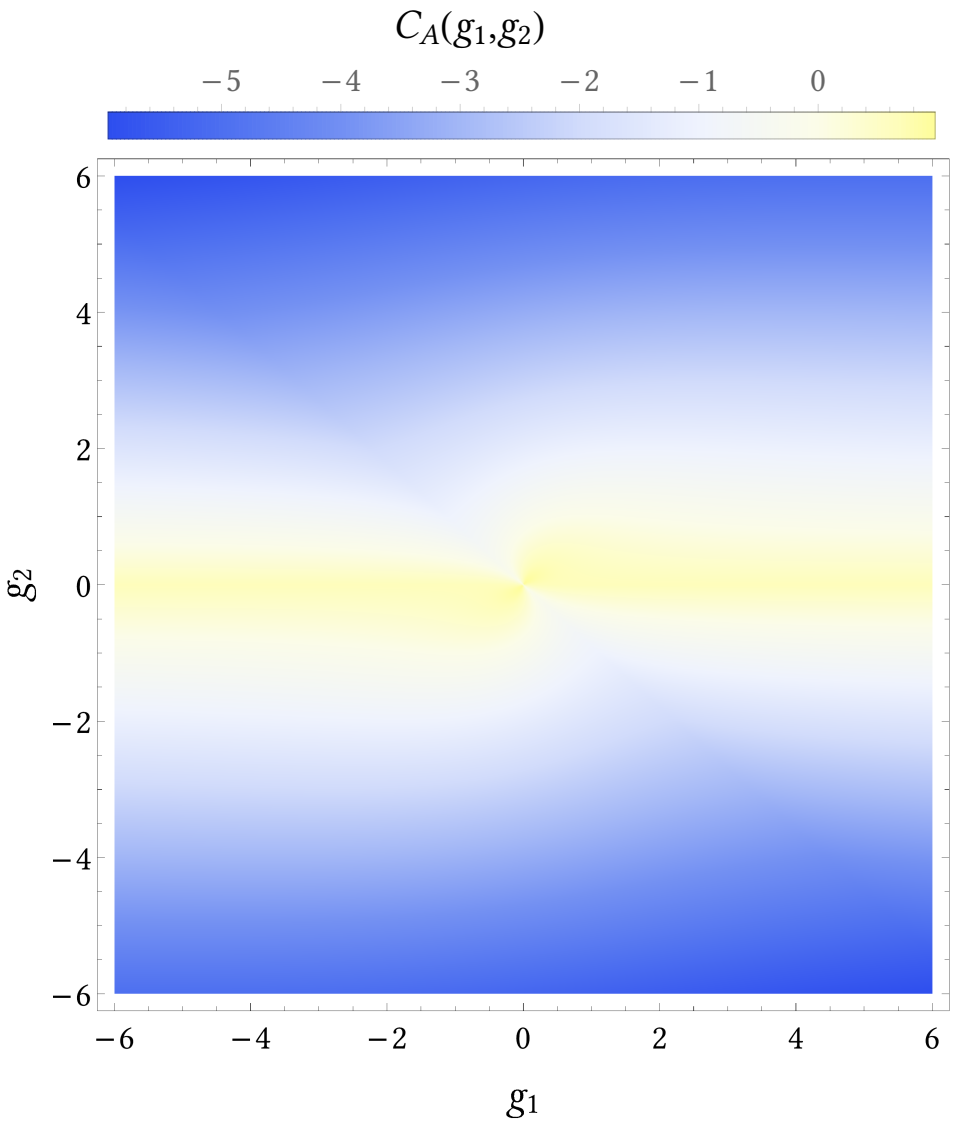



```
(*Export[
  FileNameJoin[{NotebookDirectory[],"AdamOptimizer_FirstOrderMomentsComponents.svg"}],
  plotUpdate
 ];*)
```

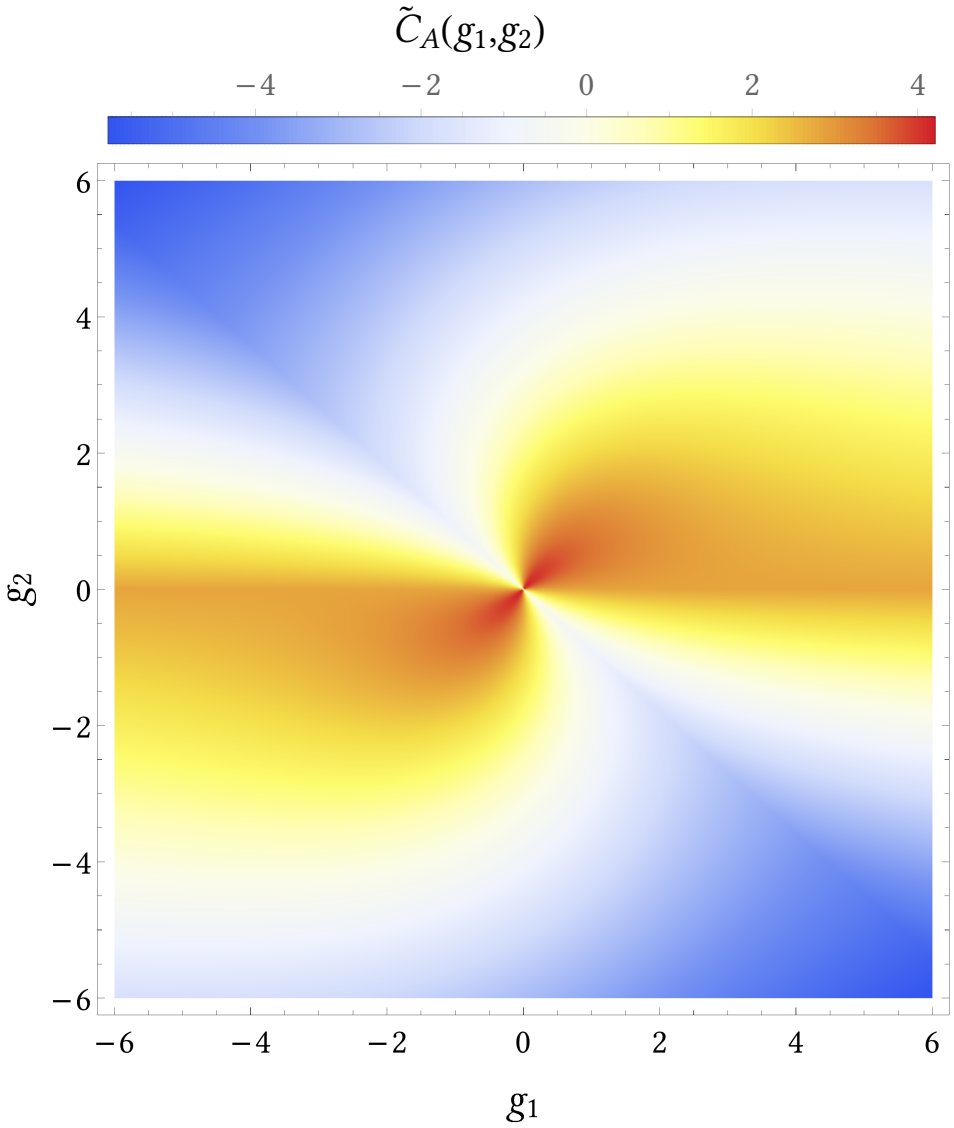
Gradient function

```
adamMeasure[g1_, g2_, beta1_: 0.9, beta2_: 0.999] := Module[{m2, s2, η = 1},
  m2 = beta1 * (1 - beta1) * g1 + (1 - beta1) * g2;
  s2 = beta2 * (1 - beta2) * g1^2 + (1 - beta2) * g2^2;

  Abs[η * (m2/(1-beta1^2))/Sqrt[s2/(1-beta2^2) + 10^-8]] - Abs[η * g2]
 ]
adamMeasureNoCorrection[g1_, g2_, beta1_: 0.9, beta2_: 0.999] := Module[{m2, s2, η = 1},
  m2 = beta1 * (1 - beta1) * g1 + (1 - beta1) * g2;
  s2 = beta2 * (1 - beta2) * g1^2 + (1 - beta2) * g2^2;

  Abs[η * m2/Sqrt[s2 + 10^-8]] - Abs[η * g2]
 ]
minValue = Min[
   Minimize[{adamMeasure[g1, g2], -6 ≤ g1 ≤ 6, -6 ≤ g2 ≤ 6}, {g1, g2}][[1]] // N,
   Minimize[{adamMeasureNoCorrection[g1, g2], -6 ≤ g1 ≤ 6, -6 ≤ g2 ≤ 6}, {g1, g2}][[1]] // N
  ];
maxValue = Max[
   Maximize[{adamMeasure[g1, g2], -6 ≤ g1 ≤ 6, -6 ≤ g2 ≤ 6}, {g1, g2}][[1]] // N,
   Maximize[{adamMeasureNoCorrection[g1, g2], -6 ≤ g1 ≤ 6, -6 ≤ g2 ≤ 6}, {g1, g2}][[1]] // N
  ];
scale[x_] := Piecewise[{{x/maxValue * 0.5 + 0.5, x ≥ 0}, {(x-minValue)/(0-minValue) * 0.5, x < 0}}];
```

```
plotAdam = DensityPlot[adamMeasure[g1, g2], {g1, -6, 6}, {g2, -6, 6},
  PlotTheme → "myTheme",
  ImageSize → 500,
  ColorFunctionScaling → False,
  ColorFunction → Function[{x}, ColorData["TemperatureMap"][scale[x]]],
  Exclusions → None,
  PlotPoints → 200,
  FrameLabel → {"g1", "g2"},
  PlotLegends → Placed[
    BarLegend[Automatic, LabelStyle → fontTicks, LegendMarkerSize → 462, LegendMargins →
      {{42, 0}, {0, 0}}, LegendLabel → Placed[Style["CA(g1,g2)", fontLabels], Above]]
    , Above]
 ]
```



```
plotAdamNoCorrection = DensityPlot[adamMeasureNoCorrection[g1, g2], {g1, -6, 6}, {g2, -6, 6},
  PlotTheme → "myTheme",
  ImageSize → 500,
  ColorFunctionScaling → False,
  ColorFunction → Function[{x}, ColorData["TemperatureMap"][scale[x]]],
  Exclusions → None,
  PlotPoints → 200,
  FrameLabel → {"g1", "g2"},
  PlotLegends → Placed[
    BarLegend[Automatic, LabelStyle → fontTicks,
     LegendMarkerSize → 462, LegendMargins → {{42, 0}, {0, 0}},
     LegendLabel → Placed[Style[Row[{Subscript["C̃", it["A"]], "(g1,g2)"}], fontLabels], Above]]
    , Above]
 ]
```



$$\text{Round}\left[\frac{\sqrt{1-0.999^2}}{1-0.9^2}, 0.0001\right]$$

0.2353

$$\left(1-0.999^2\right) * 10^{-8}$$

1.999×10^{-11}

```
(*Export[
 FileNameJoin[{NotebookDirectory[],"AdamOptimizer_BiasCorrectionEnabled.svg"}],
 rasterizeBackground[plotAdam]
];
Export[
 FileNameJoin[{NotebookDirectory[],"AdamOptimizer_BiasCorrectionDisabled.svg"}],
 rasterizeBackground[plotAdamNoCorrection]
];*)
```